INFRASTRUCTURE PLANNING FOR DROUGHT MITIGATION UNDER CLIMATE CHANGE

BY

RUIJIE ZENG

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Civil Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2012

Urbana, Illinois

Adviser:

Professor Ximing Cai

Abstract

Droughts continue to be a major natural hazard and mounting evidence of global warming confronts society with a pressing question: Will climate change aggravate the risk of drought at local scale? Are current infrastructure and their operation enough to mitigate the damage of future drought, or do we need in-advance infrastructure expansion for future drought preparedness? To address these questions, this study presents a decision support framework based on a coupled simulation and optimization model.

A quasi-physically based watershed model is established for Frenchman Creek Basin (FCB), where groundwater based irrigation plays a significant role in agriculture production and local hydrological cycle. The model, revised from the Soil and Water Assessment Tool (SWAT), simulates the dynamic response of aquifer and baseflow to groundwater pumping. The physical model is used to train a statistical surrogate model, which predicts the watershed responses under future climate conditions. The statistical model replaces the complex physical model in the simulation-optimization framework, which makes the models computationally tractable.

Decisions for drought preparedness include traditional short-term tactical measures (e.g. facility operation) and long-term or in-advance strategic measures, which usually require capital investment. A scenario based three-stage stochastic optimization model assesses the roles of strategic measures and tactical measures in drought preparedness and mitigation. Considering uncertainties involved in different climate prediction horizons, the model

results show the relative roles of mid- and long-term investments and the complementary relationships between wait-and-see decisions and here-and-now decisions on infrastructure expansion and irrigation system operations. Infrastructure expansion is preferred for the long-term plan than the mid-term plan, i.e., larger investment is proposed in 2040s than the current, due to a larger likelihood of drought in 2090s than 2040s.

Acknowledgements

This project would not have been possible without the support of my advisor Professor Ximing Cai, who provides his insight, guidance and encouragement during this research. I thank Dr.Mashor, Spencer and other group members for helpful discussions. I am grateful to the National Science Foundation (NSF) Division of Civil, Mechanical and Manufacturing Innovation (CMMI) for financial support under Award No. 0825654. I also thank my family for their mental support during my study.

Table of Contents

Chapter 1

Introduction

1.1 Background

Today, concern over drought is widespread. On average, 35-40% of the area of the United States has been affected by severe droughts in recent years [71]. Of the 46 U.S. weather-related disasters between 1980 and 1999 causing damage in excess of 1 billion, eight were droughts. Among these, the most costly national disaster was the 1988 drought, with an estimated loss of 40 billion [53]. These huge drought damages have emphasized the need to move from crisis management, which emphasizes emergency response, to risk management, which would place greater emphasis on preparedness planning and mitigation actions [70]. Traditional drought crisis management decisions only address tactical measures, i.e., post-impact responses to drought hazards under de facto infrastructure [68]. Risk management decisions include strategic measures, which are long-term or in-advance and usually require capital investment. Strategic measures can be structural, such as water storage, or nonstructural such as long-term institutional reforms for water conservation.

According to the Fourth Assessment Report (AR4) of the Intergovernmental Panel on Climate Change(IPCC) [49], "Warming of the climate system is unequivocal". One of the most important indirect issues linked to climate change relates to water supply, which is essential for most human activities, including agriculture and ecological conservation issues. Mounting evidence of global warming confronts society with a pressing question: Will climate change aggravate the risk of drought at the regional or local scale? According to AR4, droughts have become longer and more intense, and have affected larger areas since the 1970s; the

land area affected by drought is expected to increase and water resources availability in affected areas could decline as much as 30% by mid-century. In particular, U.S. crops that are already near the upper end of their temperature tolerance range or depend on heavily used water resources could suffer with further warming. Drought will revisit many areas, most probably with a greater frequency and severity than in recent memory. It is important to explore what additional risk will be imposed by climate change and what level of strategic measures should be undertaken now to avoid vulnerable situations in the future.

Unfortunately there is still much uncertainty in the climate predictions which are needed to assess drought risk [23]. The difficulty lies in the fact that general circulation models (GCMs) which have projected global climate change cannot adequately resolve factors that might influence regional climates, and also cannot tell us whether the occurrence of extreme events will increase, because they focus on long-term, large-scale averages in climate. Few studies have explicitly incorporated various uncertainties of regional climate change into drought risk estimates at the local level.

Wilhite et al. [72] suggested a holistic management framework to integrate risk and crisis management decisions and coordinate planning of strategic and tactical measures. If infrastructure capacity is not sufficient (i.e., strategic measures are limited), even the best tactical mitigation measures will not prevent large drought damage when a serious drought occurs; whereas excess infrastructure capacity means extra cost but less costly tactical measures may be sufficient to prevent a certain level of drought damage even under a limited infrastructure capacity. A quantitative tool is needed to analyze the trade-off relationships between strategic and tactical drought mitigation measures. Models have been used for analyzing either risk[30] or crisis management decision [50, 31], but not for an integrated analysis of both, and few have included the uncertainty in climate change projections and the thresholds existing in infrastructural and socioeconomic impacts.

1.2 Scope of Thesis

The Republican River Basin (RRB) is located in the high plains of northeastern Colorado and western Kansas and Nebraska. Ground water pumping for irrigation of croplands in RRB was limited prior to World War II but progressed rapidly in the 1960s and 1970s. Since then the headwater zone has been shrinking constantly in the past fifty years. For example, Frenchman Creek once originated near the town of LeRoy in Logan County, Colorado, 50-km west of Nebraska border; today the river is entirely within Nebraska. According to Burt et al. [10], a strong statistical relationship was found between the logarithm of stream flow and number of wells, current and lagged annual precipitation, and two variables that are the geometric mean of precipitation and number of wells in the current year and the year before last. Estimated mean stream flow from the statistical model in 1998 is approximately one third of that in 1950. Water rights to stream flow in the Republican River are currently in dispute between Kansas and Nebraska under a federal interstate river compact. Kansas has filed suit against Nebraska in the U. S. Supreme Court with the contention that decreases in annual estimates of virgin stream flows are not due to natural phenomena but are due to increases in groundwater withdrawals [7].

Frenchman Creek is a spring-fed waterway that begins in Phillips County, Colorado, crosses Chase and Hayes counties in Nebraska and ends at its juncture with the Republican River in Hitchcock County, Nebraska. In this study, we choose the up and middle down reach of Frenchman Creek Basin (FCB) above Ender reservoir as a case study to avoid the effect of reservoir on stream flow. The annual precipitation from 1943 to 1994 in this region is 443 mm; the precipitation increases from east to west as the effect of elevation. The annual evapotranspiration in this region is abut 1300 mm, making FCB a semi-arid region. Due to the shortage of precipitation, agriculture are heavily dependent on groundwater based irrigation. Declining groundwater levels associated with irrigation wells in the FCB appear to be associated with declining stream flow in the area, because many of the streams in the

basin receive a portion of their flow from the aquifers. As groundwater levels decline, hydraulic gradients toward the streams are reduced, thus reducing aquifer-to-stream discharge. Nearby wells can lower the water table to the point where the hydraulic gradient to the stream is reversed, thus causing stream flow depletion.

To manage the groundwater resources in this region, Nebraska water law provides a decentralized system of control to accommodate scarcity of supplies, both with respect to surface and groundwater. The state is divided into 23 natural resource districts (NRD) that have a great deal of autonomy. The regulations imposed by an NRD can include well spacing, groundwater withdrawals, rotation of pumping, and a moratorium on new wells. Since 1990s, the NRD has reduces the water permit in FCB from 20 inch to 12 inch per year. In this study we propose infrastructure expansion as a way to ease the water shortage problem is FCB. The first point is to understand the interaction of surface water and groundwater. As groundwater is a source of both stream flow (e.g. baseflow) and irrigation, human interface (agricultural pumping) is also a driven force as precipitation in FCB. With the watershed response of mid- and long-term climate forecasting, the decision making framework is to answer several questions: 1) how decisions are effected by climate change, 2) should strategic measures (e.g., infrastructure expansion) be prioritized compared to tactical measures (e.g., facility operation) for drought preparedness and mitigation under climate change, 3) what is the relative role of mid-term and long-term investments, 4) should the world invest infrastructure now or wait-and-see given the uncertainties from mid- and long-term climate change prediction?

1.3 Thesis Outline

A semi-distributed hydrological-agronomic model built in FCB is introduced in Chapter 2. The model is a modified version of Soil and Water Assessment Tool (SWAT), which

can simulate the aquifer-stream interaction under groundwater pumping. In chapter 3, climate data from regional climate models (RCM) down scaled from general circulating model (GCM) are analysis first. Due to the computation expense, a statistical-surrogate model is built from the training data from SWAT. Then the watershed responses to infrastructure expansion and facility operation are predicted by the surrogate models with different climate scenarios. In chapter4, a scenario-based three-stage stochastic optimization model is built according climate scenario tree. The decision framework provides decision maker information on how to arrange capital investment for drought preparedness and drought mitigation. Results are summarized and further work are discussed in the last chapter.

Chapter 2

Watershed Modeling

Groundwater (GW) and surface water (SW) are not isolated components of the hydrologic system, but instead interact in a variety of physiographic and climatic landscapes. Thus, development or contamination of one commonly affects the other [57]. Focusing on only one component of the hydrologic system, such as a river or an aquifer, is usually partly effective as each hydrologic component is inter-connected with other components. Understanding the interaction between groundwater and surface water is very important to hydrologists and water resources managers.

In recent years, it has been difficult to build surface water storage reservoirs along the stream due to environmental and ecological concerns. Alternatively, using an aquifer system for temporary storage of water becomes popular. In RRB, water stored in aquifers is pumped out during the growing season and recharged back to aquifers during non-growing season. Agricultural pumping has affected the processes of GW-SW interaction and gave rise to the water rights conflicts between GW-SW stake holders. Thus the conjunctive management of GW-SW leads to sustainability of hydrological, agronomic and ecological issues.

This chapter first introduces the GW-SW interaction processes under natural and aquifer pumping conditions, then describes our modeling effort to simulate the GW-SW interaction with agricultural pumping at the watershed scale.

2.1 Effect of Groundwater Pumping on Stream Flow

The interaction between streams and aquifers takes places in three different ways: streams gain water by outflow from the aquifer through the streambed (gaining streams), streams lose water by the outflow to aquifer (losing streams), or they do both, gaining in some reaches or in some seasons and losing in other reaches or in other seasons [73]. Losing streams can be hydraulically connected to groundwater by a continuous saturated zone below the stream or can be disconnected by an unsaturated zone. A special disconnected case is that the water table may have a discernible mound below the stream, when the flow rate through stream to groundwater is larger than the groundwater lateral flow rate of the mound. In this condition, groundwater pumping near the stream does not affect the stream flow near the pumping wells as the stream is suspended above the groundwater and thus not affected by aquifer.

The discharge from groundwater to stream during non-storm events sustains the stream and is vital to maintain the the fluvial ecologic communities. This part of water is usually called baseflow, which is the part of stream water coming from groundwater (some authors also include water from the soil profile). There are many hydrograph separation techniques that compute baseflow from total stream flow. Baseflow index (BFI), which is the ratio of baseflow to stream flow, describes the dependence of stream flow on groundwater. According to [10], the BFI in RRB is about 80%, even as high as 90% in FCB. In these regions, stream flow is highly dependent on groundwater, changes in groundwater storage affects surface runoff and leads to water rights conflicts between groundwater and surface water use, as we can see in RRB [61, 62].

Under natural conditions, groundwater storage reaches a dynamic equilibrium year by year. In a dry year, groundwater gets less recharge and sustains the steam flow, which decreases the groundwater storage and groundwater table. In wet years, groundwater levels recover

due to excess surface recharge. Over a long time, the groundwater storage remains stable, that is, the recharge equals discharge. In addition to the effect of climate on groundwater, human activities also affect the groundwater a lot. Humans pump groundwater for agricultural, industrial and domestic purposes. When SW and GW are directly connected, withdrawing from groundwater (especially from unconfined aquifers) can have a significant effect on the movement of water between the two water bodies. The effect of a single well or a small amount of pumping may not be significant. However, the effect of many wells continuously pumping from an aquifer may be a regional issue.

If stream and groundwater is hydraulically connected, pumping can effect the exchange water amount or even the flow direction. For a gaining stream under pre-development condition, water flow from aquifer to stream. After groundwater pump, a drawdown cone begins at the center of the well. The diameter of the cone and the drawdown develops if pumping continues. This depletes some of the water stored in the aquifer and reduces the flow to the stream. After the drawdown cone reaches the stream, the flow of stream-aquifer system changes the direction and begins to flow from stream to aquifer. At this point, the gaining stream becomes losing stream. With hydraulic connection, the rate of flow between stream and aquifer is controlled by the hydraulic gradient between the stream water surface and the groundwater table, and the conductivity of stream bed. If the pumping further continues, the groundwater table is below the streambed. The stream-aquifer becomes unconnected by a unsaturated zone. In disconnection condition, as long as the water level in stream does not change, a further drawdown of the water table due to pumping does not significantly affect the seepage from the stream. This condition can be observed when pumping does not affect the stream flow. The effect groundwater pumping on stream flow has been explored by many researchers to specify well locations and pumping schedules that would minimize the harmful effects. Several analytical solutions are available for computing drawdowns and stream depletions caused by pumping near a stream [60].

With the development of numerical groundwater models, researchers takes into account more realistic hydrologic conditions of stream aquifer systems. To assess the predictive accuracy of Glove's stream-aquifer analytical solution, which are widely used in administering water rights, Sophocleous [58] evaluate the impact of the assumed idealizations on administrative and management decisions. They compares the predictive capabilities of Glover steam-aquifer depletion model against the MODFLOW numerical standard and ranks the relative importance of the various assumptions on which the analytical model is based. The numerical model is more flexible than the analytical solution in the boundary definition and considering the spatial heterogeneity of aquifers. Chen et al. [14] further use numerical model investigate two components of stream depletion: baseflow reduction and induced stream infiltration occurring between streams and aquifers where groundwater is pumped seasonally. The former represents a pumping well capturing the base flow, which, under natural conditions, would have discharged to the stream; the latter represents the stream water recharged to the aquifer, which was induced by the pumping well. They find that baseflow reduction can be the major component in the total stream depletion.

Under predevelopment condition, the long-term groundwater system is in equilibrium state, where groundwater storage is a constant or varies about some average condition in response to annual or longer-term climatic variations. For the aquifer control volume, the water budget writes like:

$$R_{nat} = D_{nat} \tag{2.1}$$

where R_{nat} is the groundwater recharge volume under natural condition and D_{nat} is natural groundwater discharge. With groundwater pumping, human changes the natural groundwater system equilibrium. The aquifer water budget then becomes:

$$\Delta S = R_{nat} + \Delta R - D_{nat} - \Delta D - Pump \tag{2.2}$$

where ΔS is the removal of water that stored in aquifer, ΔR is the change in recharge change after human pumping, ΔD change in discharge change after human pumping, and $Pump$ is pumping volume. From equation (2.2), we see the source of groundwater withdrawal must be supplies by (1) more water entering groundwater system (increased recharge), (2) less water leaving the system (decreased discharge), (3) removal of groundwater storage, or the combination of the three. It is the changes in the system that can be withdrawn. That is, the water pumped must come from some change of flows and from removal of storage. An increases in recharge or an decrease in discharge need to balance the pumping when the groundwater system reaches a new equilibrium state after human pumping.

Between the natural equilibrium and equilibrium with pumping, there is a transient state when the system will undergo some drawdown in water levels near pumping well to induce the flow of water to the wells (forming the drawdown cone). During this process, some water initially is removed from storage. Thus, groundwater storage provides a transient source of water which gradually converts to surface-water depletion. The timing of the transition is dependent on the distance between pumping well and river, and aquifer properties, which are highly variable from case to case (e.g. months to hundred of years) [5]. To simulate the effect of human pumping on stream depletion (from the natural equilibrium to equilibrium under pumping), a model must capture the dynamic response between pumping, stream flow and aquifer storage.

2.2 Current Modeling Effort

Attempts to simulate the impact of groundwater pumping has been made by both surface hydrology and groundwater community. Emphasis has been placed on different aspects of

the natural process. Groundwater model, such as MODFLOW developed by USGS, has been applied to direct pumping regulation at watershed scale. Republican River Compact Administration Ground Water Model (RRCA Model) has been developed to determine the amount, location, and timing of stream flow depletions to the Republican River caused by well pumping and to determine stream flow accretions from recharge of water imported from the Platte River Basin into the Republican River Basin. The model solves the finite difference form of the partial differential governing equation of water balance in a discretized aquifer domain. It takes precipitation-dominated recharge, evaporation, pumping, aquifer properties as input and handles the spatial heterogeneity in a cell-by-cell way. This model also simulates the interaction between stream and aquifer. Extensive calibration on aquifer property is carried out by experts to match the groundwater head in observation wells and baseflow on some tributaries. However, the groundwater model does not treat surface water in much details. Groundwater recharge is set constantly proportional to precipitation without considering the carry-over effect of soil moisture and effect of pumping on recharge. Groundwater evaporation is reduced from potential evaporation by area-specific coefficient and groundwater pumping is estimated from energy consumption data. The simplification of surface water treats these components not as hydrologic process but as known inputs or parameters. Also, the stream flow in groundwater model is only consist of baseflow, that is, only the flow from aquifer is simulated. Surface flow (saturation excess runoff or infiltration excess runoff) which can be a big component in storm season is ignored. The stream-aquifer interaction thus is only the interaction between baseflow and aquifer. In addition, this study focus on the infrastructure planning for watershed management but the groundwater model does not provide much flexible access to represent the infrastructure in the model.

Other studies attempt to couple groundwater model with watershed model to assess the impact of irrigation impact at watershed scale [59, 16, 35]. Sophocleous et al. [59] integrate the quasi-distributed watershed model SWAT with the fully-distributed ground-water model MODFLOW to quantify the impacts of irrigated agriculture in Kansas. This inte-

gration is fulfilled by replacing the conceptual groundwater part in watershed model with physically-based distributed groundwater model. Then the groundwater model simulates the stream-aquifer interaction. This approach requires the inputs from both watershed model and groundwater model and needs converting the information in groundwater model cells into hydrological response unit in watershed model back and forth. This approach simulates all the components in hydrological processes and allows a complete analysis of the land-based hydrologic cycle, providing the means for evaluating the impacts of land use, irrigation development, and climate change on both SW and GW resources. However, this model requires inputs from both watershed model and groundwater model. Also the calibration targets include steam flow, groundwater level, and irrigation water uses. The model predictability is at the expense of model complexity, which reduces the model flexibility and increases the modeling effort. The input data uncertainty, such as steambed hydroconductivity, may overwhelm the improvement in model [24, 35].

Instead of integrating watershed model and groundwater model, this study modified the groundwater routine in SWAT to model the effects of pumping to assess the the infrastructure expansion on watershed response. The reasons to choose SWAT model, model modification, model setup, calibration and validation processes are described in the next session.

2.3 Model Description

SWAT is a conceptual, continuous time model that was developed in the early 1990s to assist water resource managers in assessing the impact of management and climate on water supplies and non-point source pollution in watersheds and large river basins [4]. SWAT is the continuation of over 30 years of model development within the US Department of Agriculture's Agricultural Research Service and was developed to scale up past

field-scale models to watershed scale. Model components include weather, hydrology, erosion/sedimentation, plant growth, nutrients, pesticides, agricultural management, stream routing and pond/reservoir routing.

SWAT delineates a watershed according to topographic data and divides the watershed into subbasins connected by a stream network. Each subbasin is further divided into hydrologic response units (HRUs), which consist of unique combinations of land cover and soil type, and there is not interaction between HRUs within a subbasin. In other word, there is no spatial differences between HRUs, which is computationally efficient by lumping similar soil and land use areas into a single unit. Computation is based on each HRUs and summed at subbasin level, then the results in subbasin level are routed though the river network. In SWAT model, groundwater is represented in SWAT by two aquifers: shallow aquifer, which is an unconfined aquifer that contributes to flow in the main channel or reach of the subbasin; and deep aquifer, which is a confined aquifer. Water that enters the deep aquifer is assumed to contribute to stream flow somewhere outside of the watershed, thus as a sink of water. In general, shallow aquifer represents the unconfined aquifer which may hydraulically connected with the stream, while deep aquifer represents the confined aquifer, which does not affect the surface water in the simulation horizon.

One important feature of SWAT model is the flexibility to represents agricultural production. SWAT model has a built-in database which describes the growth cycle of different crops. Potential/optimal plant growth, i.e. plant growth under ideal growing conditions (adequate water and nutrient supply and a favorable climate) is calculated by the method of cumulated heat degree method. This method assumes that the rate of growth is directly proportional to the increase in cumulated temperature until the number of heat to maturity is reached. Differences in growth between plant species are defined by the parameters contained in the plant growth database. The optimal growth rate is then discounted to actual growth, accounting for the varies from potential growth due to extreme temperatures, water

deficiencies and nutrient deficiencies. Agricultural management, such as drainage, irrigation, fertilizer application, tillage and rotation, can be specified by schedule.

Irrigation in an HRU may be scheduled by the modeler or automatically applied by SWAT in response to a water deficit in the soil. In addition to specifying the timing and application amount, SWAT allows to specify the source of irrigation water. Water applied to an HRU is obtained from one of five types of water sources: a reach, a reservoir, a shallow aquifer, a deep aquifer, or a source outside the watershed. Due to the lack of detailed irrigation data, SWAT can simulate auto-irrigation with a specified water stress threshold. The water stress is the difference between field capacity and soil moisture content. Once the soil water stress is below the specific value, the model will automatically apply water to the HRU. If enough water is available from the irrigation source, the model will add water to the soil until it is at field capacity. Actual plant growth equals optimal growth multiple by water stress, where water stress is unitless.

A key strength of SWAT is a flexible framework that allows the simulation of a wide variety of conservation practices and other BMPs, such as fertilizer and manure application rate and timing, cover crops (perennial grasses), filter strips, conservation tillage, irrigation management, flood prevention structures, grassed waterways, and wetlands [25]. The majority of conservation practices can be simulated in SWAT with straightforward parameter changes. Current SWAT model simulates the effect of BMPs at HRU level, so the location of BMPs are lumped in subbasin but spatially distributed in basin.

To summarize, SWAT model is chosen as simulation framework in this study as it can: (1) take future climate prediction as input to represent future climate, (2) simulate hydrological processes and crop production at the same time, (3) represent landuse management and agricultural practice in the model settings.

However, one disadvantage of surface watershed model (like SWAT) is that it does not treat ground water in much detail [59]. Particularly in this study, we want to assess the impact of groundwater irrigation while groundwater based irrigation in SWAT model is treated as having no impact on the source aquifers, thus no impact on stream flow. As discussed in 2.1, groundwater pumping breaks the aquifer natural equilibrium and reaches a new equilibrium where the change in recharge or discharge balances pumping. During the transient state, groundwater storage is depleted and provides part of pumping sources. A modification on SWAT model to relate groundwater irrigation, aquifer storage and stream flow is described below.

2.3.1 Modification on Groundwater Representation

In SWAT model, the shallow aquifer contributes base flow to the main channel or reach within the subbasin. Base flow is allowed to enter the reach only if the amount of water stored in the shallow aquifer exceeds a threshold value specified by the user, $aq_{shthr,q}$. The baseflow calculation is derived from pre-development assumption, that is, non-steady-stage water table fluctuations are only affected by groundwater recharge and baseflow:

$$Q_{gw,i} = \begin{cases} Q_{gw,i-1} exp[-\alpha_{gw}\Delta t] + w_{rchrg,sh}\left(1 - exp[-\alpha_{gw}\Delta t]\right) & aq_{sh} \geq aq_{shthr,q} \\ 0 & aq_{sh} \leq aq_{shthr,q} \end{cases} \tag{2.3}$$

where $Q_{gw,i}$ is the groundwater flow into the main channel on day i, α_{gw} is the baseflow recession constant, Δt is the time step (1 day), $w_{rchrg,sh}$ is the amount of recharge entering the shallow aquifer, aq_{sh} is the amount of water stored in the shallow aquifer and $aq_{shthr,q}$ is the threshold water level in the shallow aquifer for groundwater contribution to the main channel to occur. The baseflow recession constant, α_{gw}, is a direct index of groundwater flow response to changes in recharge. Values vary from 0.1-0.3 for land with slow response to recharge to 0.9-1.0 for land with a rapid response. Although the baseflow recession constant

may be calculated, the best estimates are obtained by analyzing measured stream flow during periods of no recharge in the watershed [74]. When the shallow aquifer receives no recharge, equation (2.3) simplifies to:

$$Q_{gw} = \begin{cases} Q_{gw,0} exp[-\alpha_{gw} \Delta t] & aq_{sh} \geq aq_{shthr,q} \\ 0 & aq_{sh} \leq aq_{shthr,q} \end{cases} \tag{2.4}$$

From equation (2.3), we can see that the baseflow in SWAT model is only dependent on the baseflow on the previous day and aquifer recharge. This formulation ignores the effect of pumping on stream depletion. The derivation in equation (2.3) only considers aquifer recharge and baseflow as the cause of water table change, while in this case, large scale groundwater pumping is reason of groundwater table decline. If the time step is large enough (large Δt in equation (2.3)), then the exponential part of baseflow (Q_{gw}) part is very small and the exponential part of recharge ($w_{rchrg,sh}$) is near 1. Then it turns out that baseflow equals recharge. If we regard baseflow as aquifer discharge, then it turns out discharge equals recharge for the aquifer. Under human development, this formulation can not represent the aquifer transient state response and new equilibrium. As discuss in section 2.1, groundwater pumping firstly obtains water from water storage in aquifer as the drawdown cone forms. The water table gradient near the drawdown cone captures water to pumping well, which will be discharged into steam without pumping. So from the equation (2.3) used in SWAT model, groundwater based irrigation is treated as having no impact on the source aquifers and stream flow.

In this study, we modify the groundwater discharge equation (2.3) to make the aquifer storage and streamflow interact with groundwater pumping. The shallow aquifer is formed as a linear reservoir, so the baseflow is linearly proportional to the shallow aquifer storage:

$$Q_{gw,i} = \alpha_{gw} S_{sh,i} \tag{2.5}$$

where Q_{gw} is the groundwater flow into the main channel at time t, α_{gw} is the baseflow recession constant, and $S_{sh,i}$ is the amount of water stored in the shallow aquifer at the beginning of day i.

Then shallow aquifer storage is updated every computation step by water budget equation:

$$S_{sh,i} = S_{sh,i-1} + w_{rchrg,sh} - Q_{gw} - w_{revap} - w_{pump,sh} \tag{2.6}$$

where $S_{sh,i}$ is the amount of water stored in the shallow aquifer, $w_{rchrg,sh}$ is the amount of recharge entering the shallow aquifer, Q_{gw} is the groundwater flow, or base flow, into the main channel , w_{revap} is the amount of water moving into the soil zone in response to water deficiencies, and $w_{pump,sh}$ is the amount of water removed from the shallow aquifer by pumping, which is calculated from in auto-irrigation commands in SWAT model.

The linear reservoir model has been widely used in lumped watershed model for a long time since Maillet [43]. When aquifer discharge is the only outflow of aquifer, the aquifer water budget is:

$$\frac{dS}{dt} = -Q_{gw} \tag{2.7}$$

Substituting equation (2.5) into euqtion (2.7) and integrating, we get the baseflow exponential recession relation, as in equation (2.4). Some studies find that the log Q_{gw} against t does not follow a straight line or the plots shows convex or concave, which implies non-linearity in aquifer release [75]. To handle this problem, some studies use two or more parallel linear reservoirs to representing components of different response time [63, 28]. This means the baseflow comes from different components which have different travel time in the watershed. Some studies find the non-linearity comes with the scale issue [75, 17]. From an experimental

watershed, Clark et al. [17] find the linear groundwater release relationship is approximately consistent with a linear reservoir at hill slope scale, and there is a deviation from linearity that becomes progressively larger with increasing spatial scale. To link the gaps between different spatial scale with linear groundwater release relationship, they provides a parallel linear reservoirs which produces both linear hill slope response and non-linear watershed scale response. In this study, we apply the linear groundwater reservoir model at each HRU. Through the river network routing, the linear groundwater discharge relation in HRUs transforms to non-linear response at the outlet of watershed.

Though equation (2.5) and equation (2.6), we relate aquifer storage, baseflow and groundwater pumping in a water balance way. Aquifer discharge is calculated from linear proportion to aquifer storage and aquifer storage is updated during every computation step considering the effect of baseflow, recharge, and pumping. Then the aquifer discharge on the next computation step is calculated from the updated aquifer storage. In pre-development condition, the groundwater system is balanced between baseflow, recharge and groundwater evapotranspiration; in transient state, aquifer storage can provide part of the water for pumping; in the new equilibrium stage, the aquifer is in a new balance between pumping, baseflow and recharge. Thus we can use the modified model to assess the impact of irrigation in RRB.

There are some assumptions lying behind equation (2.5) and equation (2.6), which may limit their application. The flow direction between stream and aquifer depends on the relative height of water table in stream and aquifer. As SWAT model does not compute groundwater table, there is no way to compare the water tables. As a solution, flow direction is only allowed from aquifer to stream, that is, the commands only simulate the gaining stream situation. It is applicable in headwater zone such as FCB where streams mainly gain water from surrounding areas, while if in some catchments where streams are losing water or both losing or gaining water equation (2.5) and equation (2.6) are not enough to capture the process. Another limitation is that in current SWAT model, calculation is based on HRUs level.

There is no interaction between HRUs while the interaction at subbasin level is realized by river network. This limits groundwater as separated bucked in each HRUs, and there is no interaction between aquifer water among calculation units. That is, there is no groundwater horizontal movement in SWAT model. According to a groundwater model built by Republican River Compact Administration (http://www.republicanrivercompact.org/), the groundwater horizontal movement is not significant in FCB compared with the groundwater storage depletion. The drawdown cone formed by pumping wells change the groundwater table height and thus change the flow pattern in aquifer. These processes is not represented in SWAT model. If interests are places on these sides, other groundwater based model should be chosen for simulation. In this study, we are more interested in water budget not the details in groundwater flow pattern and the effects agricultural practice in groundwater quantity, which is affected by climate-driven irrigation.

2.3.2 SWAT Model Setup

In this subsection, we describe data to build SWAT model in FCB, the criteria to evaluation the model, and some test scenario with SWAT model.

Watershed Delineation

In watershed boundary definition, we use a 30m (1 arc second) digital elevation model(DEM) (http://seamless.usgs.gov/ned1.php) as a main source for elevation data. Further, the eight-digit USGS hydrologic unit codes (http://water.usgs.gov/GIS/huc.html) and National Hydrography stream dataset (http://nhd.usgs.gov/data.html) also used to assist to define watershed boundary and delineate the subbasins. As there is a reservoir on Frenchman Creek, the watershed outlet is chosen at a USGS stream gage above the Enders reservoir to avoid the simulation of reservoir in the watershed, which is beyond the interest of this study. Finally, we has a watershed of about 2835 km^2 with 89 subbasins.

HRUs Definition

SWAT defines HRUs based on the combination of soil type and landuse. Soil data is from the 2006 STATSGO2 database (http://soils.usda.gov/survey/geography/statsgo/). Landuse data is from 2001 National Land Cover Data (NLCD2001) (http://www.epa.gov/mrlc/nlcd-2001.html). SWAT can define one or multi-HRUs in one subbasin. As the watershed is not much heterogamous in this study, we define one HRU in one subbasin based on the dominant soil type and landuse. The landuse data is in 2001, which may not reflect the agricultural development in the 1970s and 1980s. We use the crop data from National Agricultural Statistics Service (NASS) (http://www.nass.usda.gov/Data_and_Statistics/). The NASS dataset is summarized by county, and does not provide further spatial information. We convert the crop area from NASS to equivalent area HRUs in SWAT model. Corn plant is scheduled at 10 May and harvested at 31 October. As there is no detailed irrigation data available, we use the auto-irrigation commands in SWAT to meet the crop water requirement.

Climatic Inputs

Daily precipitation and temperature (maximun and minimum) are obtained from High Plains Regional Climate Center (HPRCC) (http://www.hprcc.unl.edu/about.php). Six weather stations lie within or near the watershed is chosen based on the continuity of time series. The stations are: Holyoke (Station ID: 054082), Enders Lake (Station ID: 252741), Imperial Municipal Airport (Station ID: 254111), Wauneta (Station ID: 259020), Sedgwick(Station ID: 259020) and Wray (Station ID: 059243).

Model Evaluation

The major hydrological budget component evaluated in this study is stream flow. As the infrastructure expansion is planed for drought mitigation, we particular focus on lowflow. The monthly stream flow observation data for calibration and validation are from USGS

National Water Information System (NWIS) (http://wdr.water.usgs.gov/nwisgmap/). The stream gage at the inlet of Ender Reservoir (FRENCHMAN CREEK NEAR IMPERIAL, siteID 06831500) is chosen to avoid the impact of reservoir. Annual crop yield data are from NASS dataset. We calibration the irrigated corn yield as it is the dominant crop and main water consumption sector in this basin. The NASS summarizes crop yield in bushel per acre, while SWAT gives yield as tons per hectare. The conversion of the units depends on the moisture content in crop. We convert 1 bushel of corn to 0.0254 ton based on literature average.

Calibration and Validation

We choose eleven parameters controlling surface runoff generation, evapotranspiration, groundwater flow, and crop growth to calibrate the model. These parameters are chosen based on previous studies that shows the sensitivity of these parameters. The parameters, their descriptions, SWAT default values, range and calibrated values are listed in table 1.

Previous studies [25] summarizes statistics criteria for evaluating the performance of SWAT model. These criteria compare the simulated value with model result, focusing on particular part of interest. In this study, we choose 2 criteria for stream flow evaluation and 1 criterion for crop yield. The monthly stream flow is calibrated using root-mean-square error (RMSE) for general data (equation (2.8)) and logarithm of RMSE, which is more focused on lowflow (equation (2.9)):

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(Q_{sim,t} - Q_{obs,t})} \tag{2.8}$$

$$LOG = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(logQ_{sim,t} - logQ_{obs,t})} \tag{2.9}$$

We use Non-dominated Sorting Genetic Algorithm II (NSGA-II) to find the optimal parameter combination. NSGA-II is a famous multi-objective optimization algorithm, which provides a objective way to characterizing multi-objective problems [20]. By identifying multiple Pareto optimal candidate solutions, NSGA-II requires a Pareto-compliant ranking method, favoring non-dominated solutions. No weight is assigned to each objective and thus no prioir information on the preference of the objectives is needed. The main advantage of NSGA-II is that they allow computation of an approximation of the entire Pareto front in a single algorithm run. The main disadvantage of evolutionary algorithms is the much computationally expensive. There is no general guide on the GA parameters. The NSGA-II population is set at 200 and generation is set at 100. The crossover, mutation and other parameters are set at their default value. The calibrated parameters are listed in table 1.

The calibration and validation are conducted from 1981 to 1985 and 1986-1990 respectively, shown in figure 2. NashSutcliffe model efficiency coefficient [48] is 0.74 and 0.72 for calibration and validation period respectively, which is regarded as good for SWAT model simulation according to the summery from Gassman et al. [25]. As we focus more on the lowflow part in the calibration, we can see that the simulated flows in the flood seasons is lower than the observed one in 1981 and higher in 1983, 1986, 1988 and 1990. All of these mismatches happen in May. Another possible cause may come from the auto-irrigation commands in SWAT model. As May is the seeding season, the irrigation requirement may not be so significant. In the model, auto-irrigation irrigates the soil until field capacity once the soil moisture stays below a threshold specified for the whole growing season. However, crop may not need so much irrigation water as they just develop. The more moisture in the soil, more saturated surface runoff produces, leading to a higher stream flow than observed.

In addition, large portion of stream flow (baseflow) is dependent on aquifer and effected by irrigation, we also compare the simulated aquifer storage with one observation well. To

see the effect of groundwater irrigation over a long time, we run the model from 1965 to 1994 with the calibrated parameter set. The simulated aquifer storage is averaged over the whole basin. The observation well data is obtained from USGS Groundwater Information (http://water.usgs.gov/ogw/gwrp/activities/fundamentaldata.html). The well (Site Number: 403220101384001) located in the mid-down reach of FCB measures the depth to groundwater table. The aquifer storage and depth to groundwater are convertible if the specific storage coefficient is known. Here we just compare the trend of the two data series as the specific storage coefficient is heterogeneous over the domain [58]. Also note that the aquifer storage is an areal average and the depth to groundwater table is a point measurement. From figure 3, we can see these two time series match quite. They all experienced a decline in from 1960s and 1970s, then remained stable afterwards. This shows the transient state from the natural equilibrium to a new equilibrium. During the transient state groundwater storage is depleted to as part of the source for irrigation. The observation data shows more variability over time than the simulated one. This is because the simulated is a spatial average over the whole basin, which smooth out the temporal variability in one point.

2.4 Summary

In this chapter, we build a watershed model which focuses on the effect of groundwater-based irrigation, especially how groundwater-based irrigation breaks the natural equilibrium of aquifer budget and the transition to a new equilibrium. The original SWAT model does not consider the impact of groundwater irrigation on aquifer storage and stream flow, that is, the model does not relate the irrigation water to the source aquifer. We modify the baseflow computation with a conceptual linear groundwater reservoir relation, where the baseflow is proportional to aquifer storage. The latter is then updated at each computation including the withdrawal of groundwater pumping. Related parameters are chosen to calibrate the stream flow and crop yield with observation data by an evolution algorithm. We also

compare the simulated aquifer storage with one observation well. The model successfully reproduces the behavior of aquifer due to groundwater pumping, which shows the validity of the revision of the model.

Chapter 3

Statistical Surrogate Model

In chapter 2, we build a physically based model (PBM) which simulates SW-GW interaction under groundwater pumping in FCB. For the infrastructure planning purpose, we are coupling a optimization model with the physical model. As the physical model is non-linear, not continuous and non-derivable, classic optimization algorithms are not applicable in this case. Evolution Algorithms (EA) is then applied to a simulation-optimization modeling framework. However, EA is very computationally expensive. GA needs to run the PBM for each evaluation, which is not reasonable for such a large scale model.

In this chapter, we build an computationally efficient data driven model (DDM) to substitute the PBM from the knowledge of Machine Learning. This DDM is built on Support Vector Machine (SVM) and trained to produce the watershed responses to infrastructure expansion under different climate change scenarios. This DDM will be latter used in the optimization framework to replace the PBM.

3.1 Regional Climate Model

The RCM simulation data sets are driven by outputs from three GCMs: 1) the U.S. Department of Energy and National Center for Atmospheric Research Parallel Climate Model (PCM) [65], 2) the Community Climate System Model, version 3 (CCSM) [18], and 3)a global atmosphere-only model (Hadley) derived from the atmospheric GCM of the Hadley Centre CGCM (Hadley) [51]. These GCMs have different climate sensitivities, that is, the

global mean temperature increase when atmosphere CO_2 concentration level doubles. PCM and CCSM have low climate sensitivity, with 2.1 and 2.2 °C respectively, while climate sensitivity of Hadley is 3.3 °C, which belong to high climate sensitivity models. Compared with all available GCMs, PCM and CCSM are at the low end and Hadley is in the upper half of the range [39].

The RCMs conducts dynamical downscaling integrations from the coarse resolution from GCMs, improving the spatial resolution from 300 km to 30 km. The downscaled RCMs provides mesoscale projections for assessing potential climate change impacts at the regional scale. The RCM that prvides the climate projection to this study is a climate extension of the fifth-generation Pennsylvania State University-Nation Center for Atmospheric Prediction (PSU-NCAR) Mesoscale Model (CMM5), version 3.3 [21]. Improvements including incorporation of more realistic surface boundary conditions and cloud cover prediction from an updated global reanalysis are made from the model of Liang et al. [40]. The simulation of CMM5 is based on the cumulus parameterization scheme, which provides superior performance in downscaling U.S.Mexico precipitation seasonal-interannual variations [41]. It has been demonstrated that CMM5 has considerable downscaling skill over the United States, producing more realistic regional details and overall smaller biases than the driving reanalyses or GCM simulations [42].

The historical simulation corresponds to the coupled model intercomparison project 20th Century Climate in Coupled Models scenario (20C3M) [19], driven by historically accurate forcings, including anthropogenic emissions of greenhouse gases and aerosols, indirect effects on atmospheric water vapor and ozone, and natural changes in solar radiation and volcanic emissions. The baseline simulations for RCM derived from PCM, CCSM3, and Hadley are 1991-2000, 1990-1999, and 1980-1989, respectively, which have been used as the baseline periods for climate change impact assessment [2]. The future simulation is forced by Emission Scenarios from the IPCC Special Report [47]. Each RCM simulation scenario has 10-year

length climate series. The simulation period is 2040-2049 and 2091-2100 for PCM, 2041-2050 and 2090-2099 for Hadley, 2090-2099 for CCSM, respectively. The two PCM emission scenarios are A1Fi (high, effective CO_2 concentration of ∼970 ppm by 2100) and B1 (low, ∼550 ppm by 2100), respectively; the two CCSM emission scenarios are A1Fi and A1B (middle, ∼720 ppm by 2100), respectively; the two Hadley emission scenarios are A2 (moderately high, ∼860 ppm by 2100) and B2 (moderately low, ∼620 ppm by 2100), respectively. The RCMs simulation results are 3-hour continuous time series including precipitation, maximum and minimal temperature, wind speed, humidity and solar radiation. The 3-hour data are aggregated into daily data to feed the PBM developed in Chapter 2 to assess the climate change on watershed management.

The historical precipitation and temperature data from 1985-1994 are used to assess the climate change predicted by each prediction, shown in figure 4. The 10-year-mean annual precipitation from historical data is 519.4 mm. RCMs scenarios predict annual propitiation differently, ranging from 305.5 mm to 594.6 mm. Among the baseline scenarios, Hadley RCM has higher annual precipitation at 553 mm, CCSM RCM has lower annual precipitation at 364 mm, and PCM RCM predicts unchanged precipitation. For the future scenarios, Hadley-B2-2090s has an increasing precipitation at 595mm, Hadley-A2-2090s has unchanged precipitation, and other scenarios all experience decreasing precipitation. Among those, CCSM-A1B-2090s has the lowest precipitation, 41 % decrease from current. Also, the precipitation predicted by different GCM driven RCMs shows some patterns regardless of their scenarios. CCSM driven RCMs have the lowest precipitation, PCM driven RCMs have slightly decreasing precipitation, and Hadley has unchanged or increasing precipitation.

Although annual precipitation change shows a general trend of natural water available in the future, the monthly precipitation variation is more important for the water resources management, as the water use and hydrological regime shows a strong seasonal pattern in this watershed. The precipitation in growing seasons (May to September), shown in figure5

and figure 6, affects the complementary irrigation that would be pumped out for crop production; the precipitation in dry seasons is important for in-stream ecological community and drought management. The monthly precipitation predicted by RCMs shows great variations compared with annual precipitation. Although precipitation change in other months during 2040s is different in each model predictions, all model show increasing precipitation in June. As June is the month when evaporation is highest, the increases in precipitation can bring more water to sustain the stream flow. A1Fi scenario in PCM has decreasing precipitation during most of the month, while B1 scenario has increasing precipitation in the first half of the year and decreasing precipitation in later months of the year. For the period in 2090s, CCSM driven RCMs predicts increasing June precipitation while other monthly has decreasing precipitation. The decrease is near 80% in from July to October in A1B and A1Fi scenario. August is also the month when crop water demand is very high. The decreasing August precipitation implies the agriculture may more dependent on irrigation. The Hadley driven RCMs all have increasing spring precipitation and a slightly (within 20%) decrease in other months. The increases in spring are 100%, 50%, and 77% for baseline, A1B, and A1F respectively. The PCM driven RCMs shows more variability in monthly precipitation. The baseline has 75% increase on February, April and June; A1Fi scenario predicts decreasing precipitation, especially in July and Junel; and B1 scenarios has increase precipitation in spring and June, while 15% to 35% decrease in other months.

From crop production, precipitation in growing season is of more importance as it provides water resources for evapotranspiration. Given the potential evaporation unchanged, a decreasing precipitation in growing season usually implies an increase in complementary irrigation demand. The growing season (from May to August) precipitation during 2040s decreases 28.6% and 14% in PCM-A1Fi and Hadley-A1 respectively, and remains unchanged in PCM-B1 scenario. The growing season precipitation by CCSM model during 2090s decreases about 30%, even nearly 40% in A1B and A1F scenarios. Growing season precipitation during 2090s increases 23.3% in Hadley-B2 scenario and decreases 18.9% in PCM-A1Fi scenario.

Other scenarios have less 10% change in growing season precipitation. Generally the future climate has less precipitation as predicted by the RCMs, which implies a more stress for irrigation. Also the less precipitation indicates more likely to have drought, which requires the potential for infrastructure expansion in this region to sustain the low flow given the already depleted stream flow.

The mean temperature is averaged from the maximum and minimal temperature. Mean temperature increase in nearly all month by all the prediction, but the variation is different from scenarios. Monthly temperature change during 2040s has a slightly increase, mostly within 5 °C. For the 2090s periods, PCM-B1 scenarios has temperature increases within 5 °C, while other scenarios experiences a month temperature increase over 5 degrees. CCSM-A1F predicts a 10 degree monthly temperature increase. The increase in temperature has two impacts on the watershed: 1) high temperature implies a high evapotranspiration rate in the watershed, which may lead to a decrease in stream flow and higher demand for irrigation, 2) heat wave by unusual high temperature also stresses the crop growth even there is enough water supply.

3.2 Overview of Support Vector Machine

Hydrological models try to represent the hydrological system in mathematical form. Within a boundary, the system consists of different components and a set of equations linking the input and output variables. The variable may be function of time or space or both. Depending on the processes and details described, hydrological models can be divided into physical or scale models, conceptual models, lamped models, physically based models or empirical models [15]. Due to developments in area of machine learning, empirical models received an boost during last decades. Data-driven model is based on the analysis of all the data characterizing the system under study. A model describes the connections between the system state variables (input, internal and output variables) under a certain assumptions about the

processes within the system.

A model can then be defined on the basis of connections between the system state variables (input, internal and output variables) with only a limited number of assumptions about the physical behavior of the system. The methods used nowadays can go much further than the ones used in conventional empirical modeling: they allow for solving numerical prediction problems, reconstructing highly non-linear functions, performing classification, grouping of data and building rule-based systems. DDM is a modeling approach which focuses on using the machine learning techniques in building models of physical processes. These models can complement or replace the knowledge-driven models describing behavior of physical systems.

Support vector machine (SVM) is a concept in statistics and computer science for a set of related supervised learning methods that analyze data and recognize patterns, used for classification and regression analysis. Unlike other machine learning techniques, SVM has some unique properties which leads to its popularity: 1) good generalization performance, this is because SVM seeks to minimize the upper bound of generalization error rather than minimizing the training error, 2) the solution of SVM is always globally optimal, while many other machine learning techniques are subjected to local minimal (e.g., ANNs [76]), 3) the solution is represented by Support Vectors, which are typically small subset of all training data.

A brief overview of ϵ-SVR is provided here. For more details, readers are referred to [56] and [64]. The goal of ϵ-SVR is to find a function $f(x)$ that has at most ϵ deviation from the actually obtained targets y_i for all the training data, and at the same time is as flat as possible. That is, only errors larger than ϵ are considered. Given training data $\{(x_1, \epsilon_1) \ldots, (x_N, \epsilon_N)\} \subset \mathcal{X} \times \mathbb{R}$, where $\mathcal{X}$ denotes the space of input patterns (e.g. $\mathcal{X} = \mathbb{R}^d$). The input x_i is first projected to a higher dimensional feature space by the map $\Phi : \mathcal{X} \to \mathcal{F}$. Linear regression to approximate the unknown function $\epsilon(x)$ is then performed in the feature space $\Phi(x) = \mathcal{F}$ instead of the input space $\mathcal{X}$:

$$f(\mathbf{x}) = w \cdot \Phi(\mathbf{x}) + b. \tag{3.1}$$

The coefficients w and b are estimated by solving the optimization problem formulation of SVR:

$$minimize \ \ \frac{1}{2}||w||^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*) \tag{3.2}$$

subject to:

$$(w^T\phi(\mathbf{x}_i) + b) - \epsilon_i \leq \varepsilon + \xi_i, \tag{3.3a}$$

$$\epsilon_i - (w^T\phi(\mathbf{x}_i) + b) \leq \varepsilon + \xi_i^*, \tag{3.3b}$$

$$\xi_i, \xi_i^* \geq 0, \ i = 1, ..., N. \tag{3.3c}$$

Regularization by minimizing $||w||^2$ ensures the flatness of the solution. The second term in Eqn. (3.2) is derived from the ε-insensitive loss function:

$$|\epsilon_i - f(\mathbf{x}_i)|_\varepsilon = max\{0, |\epsilon_i - f(\mathbf{x}_i)| - \varepsilon\}. \tag{3.4}$$

The constant C in Equation (3.2) determines the trade-off between the flatness of f and the tolerance of deviation larger than ε.

Usually the map $\Phi : \mathcal{X} \to \mathcal{F}$ is implemented implicitly via kernels such that:

$$\langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \rangle = K(\mathbf{x}_i, \mathbf{x}_j) \tag{3.5}$$

where $\langle \cdot, \cdot \rangle$ denotes the dot product in $\mathcal{F}$. In this work, the popular *radial basis function* (RBF) is used as kernel:

$$K(\mathbf{x}_i, \mathbf{x}_j) = exp(-\gamma ||\mathbf{x}_i - \mathbf{x}_j||^2). \tag{3.6}$$

The kernal width parameter γ was optimized via cross validation.

Following the recommendation of Cherkassky and Ma [13], the regularization hyperparameter C and error insensitive hyperparameter ϵ are given by

$$C = max(|\mu + 3\sigma|, |\mu - 3\sigma|), \tag{3.7a}$$

$$\varepsilon = \tau \sigma_0 \sqrt{\frac{ln\ N}{N}}, \tag{3.7b}$$

where μ and σ are the mean and standard deviation of the training outputs ϵ_i's, τ is a coefficient usually which equals to 3, σ_0 denotes the noise level, and N is the size of training dataset. The reason that cross validation is not used to optimize all three hyperparameters is due to the prehibitively long computation time required by sufficiently fine grid search. In addition, the hyperparameters chosen analytically outperform the values that chosen by a preliminary cross validation attempt with a relatively coarse grid search. The codes can be downloaded from the LIBSVM website (http://www.csie.ntu.edu.tw/ cjlin/libsvm/).

3.3 BMPs Representation

The substitutionary model is built based on the simulation data of PBM with different future climate predictions and planning decisions response. In this study, the planning decisions includes strategic and tactical measures. Unlike the traditional crisis drought management, which emphasizes tactical measures for emergency response, we also introduce strategic measures as strategic measures as preparedness planning and mitigation actions [71]. Tradi-

tional drought crisis management decisions only address tactical measures, i.e., post-impact responses to drought hazards under de facto infrastructure [68]. Risk management decisions include strategic measures, which are long-term or in-advance and usually require capital investment. Strategic measures can be structural, such as water storage, or nonstructural such as long-term institutional reforms for water conservation.

The tactic measures are taken only after a certain climate happens and are operational measures to mitigate the damage caused by drought. The tactic measures include irrigation from shallow and deep pumping wells. On one hand, pumping from aquifer provides the water for crop evapotranspiration, which is not enough from precipitation; on the other hand, pumping from shallow aquifer can cause stream depletion and pumping from deep aquifer is more expensive and the effect of groundwater depletion needs more time to recover.

The strategic measures in this study are two types of BMPs: infiltration pond and terraces. Terraces are earth embankments and channels constructed across the slope at suitable spacings and with acceptable grades for one or more of the purposes to: 1) reduce soil erosion, 2) provide for maximum retention of moisture for crop use, 3) improve water quality, 4) reduce peak runoff rates to installations downstream, 5) reduce sediment content in runoff water. Terraces can be identified by the way in which they handle runoff. There are three types of terraces: 1) storage terrace, where the embankment of a storage terrace can include sections that store runoff as well as sections that intercept runoff and convey it to the storage. The water can be stored for a period of time and then released by underground outlet, 2) gradient terrace, where runoff is intercepted by the terrace and conveyed directly (without any storage) to a stable surface outlet at non-erosive velocities, 3) level terrace, where runoff is intercepted for the purpose of moisture conservation.

Infiltration pond is another type of BMPs that used to regulate stormwater runoff. Unlike detention pond, which is designed to discharge to a downstream water body, or a reten-

tion pond, which is designed to include a permanent pool of water, an infiltration pond is designed to infiltrate stormwater though permeable soil into aquifer system. Although it does not discharge to a surface water body under most storm conditions, it overflows during flood conditions. For the location of infiltration pond, the site should be a groundwater recharge zone which has uncompacted and permeable soils. The pond should fail if it is not properly maintained, especially in watershed with high levels of sediment which forms an impermeable layer at the bottom of the pond. Important benefits of groundwater infiltration facilities include reducing surface-runoff volume, reducing pollutant discharge, reducing thermal impacts on fisheries, increasing groundwater recharge, and augmenting lowflow stream conditions [45]. The purpose of building infiltration pond in FCB is to convert the stormwater into artificial groundwater recharge. The increasing recharge is a compensate for the extensive exploitation of groundwater. Groundwater recovered from the extra recharge thus sustains the lowflow during drought conditions.

The effectiveness of these BMPs have been studied with SWAT model by many researchers [9, 33, 55, 54]. Bracmort et al. [9] evaluate the long term performance of BMPs by estimating the condition of a BMPs based on visual inspection and compare them with selected original design dimensions. Although they found economic benefits received from the BMPs did not outweigh the costs for implementing and maintaining the BMPs, they claimed that these BMPs has lots of benefits pertaining to water quality, wildlife habitat improvement, human and aquatic ecosystem health, downstream impacts, and intangible impacts, which are not able to assign a market value. Kaini et al. [33] used an evolution algorithm to optimize location and size of BMPs combinations that cost effectively to promote achievement of treatment goals at large spatial scales.

As there are high uncertainties in the degradation of BMPs, we only consider the size of BMPs in this study. That is, we assume the BMPs are properly maintained which can ensure the performance as designed. If further information on the degradation, the cost associated

with maintenance and the degradation rate can be implemented into the model. Also, we do not consider the location of the BMPs and assign uniform values to each subbasin in FCB. This can be incorporated into our model but will increase the dimension of optimization problem. In this study, we focus on the benefits and costs of the BMPs over the whole basin, the improvement of considering the spatial variation of BMPs is beyond our scope.

3.4 Substitutionary Model

The substitutionary model is built by SVR from training data run from the PBM. The decision variables representing strategic and tactic measures are: 1) pond, the size of infiltration pond in each subbasin, 2) terrace, the terrace built to reduce the slope in each basin, 3) auto irrigation trigger, represented by the soil moisture below the field capacity, and 4) irrigation water source, the threshold of shallow aquifer water depth when transferring pumping from shallow aquifer to deep aquifer. In this study, due to the computation expense, we do not consider the spatial variation of each variables and apply them uniformly to each subbasin. The outputs from PBM are: 1) 10-day-flow from 10-year length simulation, 2) mean annual crop yield, 3) mean annual irrigation from shallow aquifer, and 4) mean annual irrigation from deep aquifer.

The procedures of building SVR watershed model representing response of strategic and tactic measures under different climate scenarios are as follows: 1) generating combinations of decision variables. As SVR dose not work well on extrapolation, all the decision variables are sampled from a big enough range. Considering the computation time and goodness of representation, 2500 combinations are generated from the the four decision variables, uniformly randomly sampled from their range. 2) simulating the corresponding outputs from the decision variable combinations from PBM. Each combination of decision variables is feed into PBM model with a climate scenarios, and calculate the 4 output variables. After

running all the combinations of decision variables, SVR can represent watershed response of strategic and tactic measures under a certain climate scenario from PBM. 3) repeating process 2) to build other SVRs, which represents the response of a certain climate scenario.

To handle the complexity and goodness of fitting SVR model, we adjust two parameters in SVR model: α, which represents the penalty for model complexity and σ^2, which measures the goodness of fit of model. The 2500 combinations of data sets obtained from PBM are split into 3 subset: 1800 sets for SVR training; 400 sets for α and σ^2 tuning; and 300 sets for validation. The two parameters are adjusted by maximizing the mean value of of R-squared values of the 4 outputs. As R squared value is between the range of 0 to 1, there is no necessary to normalized the value between the outputs. Thus the mean value of R-squared means the same weight on 4 outputs. The parameter values and the R-squared values during calibration and validation are listed in table 2. The SVR model has a great improvement on computation efficiency.

3.5 Catchment Response to BMPs Expansion and Irrigation Operation under Different Climate Scenarios

3.5.1 Effect of infiltration pond on low flow

Infiltration pond is designed to retain storm water and increase aquifer recharge, while can recovers aquifer storage and sustain low flow. Water captured in infiltration pond either recharges aquifer or evaporates into atmosphere. Infiltration pond with large capacity captures more storm water potential for recharge. On the other hand, large pond has a large surface area for evaporation. Figure 5 shows the effect of infiltration pond size on 10-day-flow. There shows two patterns in the response curve. Watershed low flow increases first and then decreases or remains with increasing infiltration pond size in CCSM model scenarios. In other climate scenarios, low flow increases with infiltration pond size. In the analysis of

future climate data, we have found CCSM model scenarios predict less precipitation and higher temperature. In these scenarios, when pond capacity is small, infiltration pond do not capture large amount of water, which infiltrates into the aquifer quickly. Although evaporation exists, the amount of evaporation is not dominant compared with filtration; when pond capacity increases, more water stored in the pond. At the same time, ponds with larger surface area evaporate more water. When evaporation is very high, evaporation becomes dominant. Instead of increasing recharge, the building of pond has the effect of evaporating water back into atmosphere, which would become recharges or stream flow with no or small ponds. Also, in Hadcm3-A1 scenarios during 2040s, the low flow increases with infiltration pond size at first. After the pond size reaches 18 million m3, the low flow does not increase with pond size but becomes stable or slightly decreases. This can also be explained by high evaporation rate in this scenario. Comparing precipitation in PCM-A1Fi and Hadcm3-A1 during 2040s, although the former has more decreases in precipitation than the latter (24% vs. 16%), Hadcm3-A1 has a higher mean temperature. The increases in mean annual temperature predicted are 1.9 and 4.6 by PCM-A1Fi and Hadcm3-A1, respectively. The increase in temperature indicates a higher evaporation rate. The more increases in evaporation by Hadcm3-A1 outbid the more decreases in precipitation in PCM-A1Fi, so the low flow does not increases with pond size in the former scenario, while the latter scenario can have an increasing lowflow-pond size relation. To summarize, infiltration pond captures storm water and recharges aquifer, thus increases the baseflow. The choosing of pond size is dependent on the effect of increases in low flow and the cost of pond building. However, in scenarios with extreme high temperature, the evaporation from ponds outbid the effect of their benefit. In these conditions, the climate also becomes a physical constraint for the choosing of pond size.

3.5.2 Effect of terrace expansion on 10-day-flow

The terrace has the effect of reduce sub-basin hill slope, which reduces the peak flow in storm events and increases the travel time of water in subsurface. The lagged release of subsurface

flow increases the aquifer recharge and also sustains the low flow during between-storm periods. Although the terrace increases lowflow, the effect is not as significant as compared with infiltration pond. This is because the terrace does not directly change the flow rate among different water components but changes the travel time of subsurface water and redistributes this part of water. Unlike the different patterns in lowflow response to pond size with different climate scenarios, the building of terrace increases the lowflow in all scenarios. Terrace is not affected by evaporation, thus can be a good choice of BMPs in high evaporation scenarios.

3.5.3 Effect of BMPs on crop yield

The expansion of infiltration pond or terrace does not quite affect the crop yield under the climate scenarios. This is understandable as the BMPs are mainly designed to groundwater recovery and environmental flow requirement, while crop yield is high deepened on the water available for crop evapotranspiration either from effect rainfall or complementary irrigation. The expansion of BMPs can indirectly reduce the irrigation cost as the BMPs recovers groundwater thus reduces the pumping cost. Attention should be paid to the crop yield in CCSM-A1F scenarios during 2090s. The crop yield is about 6 ton/ha in CCSM-A1F, compared with 10 ton/ha in other scenarios. The low crop yield in this scenario can be explained by the high temperature. The optimal growth rate of crop is calculated by accumulated heat units on each day. Then the actual growth rate is reduced due to water stress or temperature stress from optimal growth rate. Daily temperature deviated from the optimal temperature results temperature stress for the crop, which results an exponential decay during the actual crop yield calculation. The CCSM-A1F scenario predicts a 9.2 ^{o}C increases in daily temperature, the highest among all the scenarios. Heat wave is the dominant reason for the decrease in crop yield. The effect of high temperature on crop yield is beyond the scope of this study, but can be further studied to predict the impact of climate change on crop production.

The CO_2 concentration rate in the atmosphere is one of the assumptions made for each climate scenarios. In this study, we do not consider the CO_2 concentration rate on the crop yield but only consider the climate variables on the crop water use and temperature stress.

3.5.4 Effect of Irrigation on crop yield

Complementary irrigation supplements the water shortage of insufficient effective rainfall. The actual crop yield is reduced from the optimal growth rate when there is water stress. Larger irrigation applied in the crop increases the crop yield in all climate scenarios. The slope of Yield-Irrigation relation is big at low irrigation depth and small at high irrigation depth, which means the potential for stress irrigation. A reduce of irrigation amount around the optimal growth rate does not decrease the yield much, but can save a large amount of irrigation cost. Irrigation amount to get a near-optimal crop yield is less than 400mm in most scenarios, but is nearly 600mm in scenarios with low precipitation and high temperature, such as CCSM-A1B and CCSM-A1F.

3.5.5 Effect of irrigation from upper aquifer on low flow

As stream is hydraulically connected with aquifer, pumping from aquifer for agricultural irrigation also affects the stream flow. Figure 10 shows the effect of irrigation amount from upper on the lowflow. Generally, the groundwater based irrigation decreases the low flow due to the stream depletion effect. At high irrigation depth, low flow does not further decreases with irrigation (such as Hadley-B2, PCM-B1, PCM-A1Fi), even increases a bit (such as CCSM-baseline and CCSM-A1B scenarios). Although pumping causes stream depletion, on the other hand, large amount of pumping maintains the soil moisture close to the field capacity. High soil moisture content generates more subsurface flow and produces more saturated surface runoff. In these dry scenarios, the effect is more significant as we can see in CCSM-baseline and CCSM-A1B scenarios.

3.5.6 Effect of irrigation from lower aquifer on low flow

Lower aquifer is hydraulically unconnected with stream flow. Pumping from lower aquifer acts like introducing water from outside of this system, so the low flow all increases with pumping from lower aquifer. Pumping from lower aquifer costs more than that from shallow aquifer. Besides, pumping from lower aquifer requires more time to recover, and can causes problem outside the basin if the lower aquifer within a regional aquifer system. Pumping from lower aquifer can be used as an emergency water supply for both agricultural production and environmental conservation during very dry condition.

3.6 Summary

This chapter is a bridge which connects the simulation model in chapter 2 and optimization model in chapter 4. 12 climate scenarios of RCMs driven by different GCMs models are obtained to represent the future climate condition. The GCMs have different climate sensitivity to climate change, and RCMs have different assumptions on CO_2 concentration in 2100. The annual precipitation in the future ranges from 305.5 mm to 594.6 mm, compared with 519.4 mm in current value. 2 out of 12 predict increasing annual precipitation and 2 out of 12 predict unchanged annual precipitation, while the remaining predict decreasing annual precipitation. The lowest precipitation are predicted by CCSM based models. The monthly and growing season precipitation change show the variation during a year. Hadley model predicts precipitation increase in spring and decrease in other seasons. CCSM model predicts decreasing precipitation in all months except June. The decreasing growing season precipitation indicates more irrigation need for agriculture production. For the temperature, all models predict a increase, ranging from 0.6 to 9.2 ^{o}C. The high temperature increases the crop evapotranspiration, and the heat wave caused by extreme high temperature can

also reduce crop yield if even there is enough water.

A substitutionary model is built with SVR from the training data from SWAT model for each climate scenario. The SVR has a great improvement on computation efficiency and still represents the relation produced by PBM. The SVR simulation predicts that under future climate scenarios, infiltration pond and terrace are effective BMPs to increases the lowflow during drought conditions. In extreme low precipitation scenarios, the selection of infiltration pond size depends on the cost-effect evaluation and also physical constraints, as infiltration pond can also evaporate storm water back into atmosphere. The crop yield is less affected by BMPs expansion, but are more related to the irrigation operation. Crop yield is increased by irrigation and constrained by temperature stress as heat wave in extreme temperature can reduce crop yield. Irrigation helps to increase crop yield, on the other hand, irrigation from shallow aquifer decreases lowflow in most of the cases. In dry climate scenarios, effect of irrigation from shallow aquifer first decreases lowflow and then maintain or even increases lowflow. This is because the large amount of irrigation water applied on crop land maintains the soil moisture at field capacity. This increases from subsurface flow and excess saturated runoff by precipitation compensate the baseflow depletion caused by pumping. Irrigation from deep aquifer increases lowflow as deep aquifer is hydraulically unconnected with stream. Pumping from deep aquifer costs more and can affect groundwater flow pattern of a large region. Deep aquifer irrigation can be good source of water for drought mitigation during extreme dry condition when surface water is not enough and shallow aquifer pumping is not preferred for the consideration of lowflow conservation.

Chapter 4

Decision Making Framework for Infrastructure Expansion

In this chapter, we build a decision framework to answer those question: 1) how investment is allocated between infrastructure expansion (strategic measures) and facility operation (tactical measures) for the preparedness and mitigation of drought under an uncertain future? 2) How does the temporal allocation of investment depend on climate change projection? That is, should the world invest infrastructure now or wait-and-see given the uncertainties from mid- and long term climate change prediction? 3) How uncertainties from different climate forecasting horizons are handled in the decision framework?

4.1 Scenario Tree

Multistage stochastic programs are effective for solving long-term planning problems under uncertainty. In this study, a three stage stochastic optimization is developed for decision-making on infrastructure expansion for both short-term and long-term plan. The optimization is based on a scenarios model consist of many future climate predictions by RCM-GCM. The concept of scenarios is usually employed for the modeling of randomness in stochastic programming models, in which decisions have to be made independent upon knowing the actual paths where data evolve over time [77]. Such data are usually subject to uncertainty or some kind of risk. In this problem, the decision variables are strategic and tactical measures taken for drought mitigation and the investment decisions must be made before their performance can be assessed. Each scenario can be viewed as one realization of an underlying stochastic data process. A good approximation of the underlying stochastic process

may involve a very large number of scenarios and their probabilities. A better accuracy of uncertainties is described when scenarios are constructed via a simulated path structure [29].

The scenario generation is a process which involves : 1) modeling of randomness, which employees the set of available past information with the aim of building sub-models for each individual stochastic parameter, 2) generating a set of scenarios that encapsulate the consistent depictions of pathways to possible futures based on assumptions about economic and technological developments from the sub-models, 3) the factors driving the risky events are approximated by a discrete set of scenarios, or sequence of events [52]. According to the complexity of stochastic model, the scenario tree structure is used to approximate the random process (Heitsch and Rmisch 2005).

Due to the solvability of stochastic program, the scenario tree should be constructed from the available climate scenarios. The strategy is to generate a 3 stage scenario tree from a set of individual climate scenarios by bundling scenarios based on their property (e.g. RCM-GCM model performance or model assumption). In this problem, 2 decisions are made in the first stage; 4 and 2 decisions are made for each scenario in the second stage and third stage, respectively. A good representation of the scenarios can reduce the number of decision variable and accurate approximate of the future. The scenario tree reflects the inter-stage dependency and the decreases the number of links between scenarios.

Due to the availability of the scenarios in different stages, the scenario tree for this 3-stage optimization problem is built based on the CO_2 concentration assumptions made for each emission scenarios. The emission scenarios we have are A1Fi, A2, A1B, B1, and B2, with effective CO_2 concentration of $\sim$970 ppm, $\sim$860 ppm, $\sim$720 ppm,$\sim$ 550ppm and $\sim$ 620ppm by the year of 2100, respectively. The five scenarios correspond to high, moderately high, middle and moderately low assumptions on CO_2 concentration. The baseline scenario are for the simulation of historical climate thus represent the current CO_2 concentration level.

To make use of the baseline scenarios, we treat them as representation of 'unchanged future climate'. We can cluster the scenarios into 4 categories on future CO_2 concentration level: high (A1Fi), middle (A2 and B2)low (B1 and A1B) and unchanged. Furthermore, we assume the changing trend of CO_2 concentration level is consistent in the next century: high CO_2 concentration level in 2040s leads to high CO_2 concentration level in 2090s; low in 2040s leads to low in 2090s; and unchanged remains unchanged for the whole century. This clustering based on CO_2 concentration level forms the nodes and links in the scenario tress.

As we don't have scenarios on unchanged CO_2 concentration in 2040s, historical data from 1985-1994 are used to represent unchanged climate as baseline scenario in 2040s. The current can lead to four possible climate in 2040s: 1) unchanged (baseline), 2) high CO_2 concentration level (PCM-A1Fi), 3) middle CO_2 concentration level (Hadley-A2), 4) low CO_2 concentration level (PCM-B1). Then the unfold climate in 2040s further leads to climate scenarios in 2090s by different RCM-GCM, shown in figure 16.

One of the advantages on this scenario tree is that by bundling the scenarios by CO_2 concentration level, we average the uncertainty of different RCM-GCM models. As the RCM-GCMs have different climate sensitivity, scenarios in in 2040s lead scenarios with similar CO_2 concentration level but from different RCM-GCM model, thus averaging the uncertainty in GCM models. Another advantage is that the scenarios tree has only 36 decision variable (2 for first stage, 4*4=16 for second stage, and 2*9=18 for third stage). If the scenarios is developed as scenario fan (each scenario in second stage leads to every scenario in third stage), we would have 90 decision variables(2 for first stage, 4*4=16 for second stage, and 2*4*9=72 for third stage). This reduces the computation expense of the optimization. If the GA population is set as 5 times of the decision variables, the the problem with scenario tree needs 36*5*9=1620 evaluations and scenario fan needs 90*5*81=36450 evaluations for generation.This greatly reduces the dimension of optimization problem

4.2 Three-stage Stochastic Optimization

Stochastic programming is effective for problems where an analysis of policy scenarios is desired and uncertainties are expressed as random variables with known probability distributions. The fundamental idea behind stochastic programming is the concept of recourse [8]. Recourse is the ability to take corrective action after a random event has taken place. For example, in two-stage stochastic programming support the here and now decision, while providing a number of wait and see solutions dependent upon which scenario unfolds: a decision is firstly undertaken before values of random variables are disclosed and, then, after the random events have occurred and their values are known, a recourse action is made in order to minimize penalties that may appear due to any infeasibilities [31].

There are many studies which implement two-stage stochastic program in water resources planning and management [44, 38, 37, 67, 12, 69]. Wilchfort et al. applied a two-stage model to analyze the planning of long-term water conservation measures and the implementation of short-term conservation measures, considering economic costs of those long-term and short-term measures and probability distribution of water shortage. Cai et al. [12] extends the two-stage decision process with risk aversion analysis in the choice of irrigation technology, which provides a procedure for trade-off analysis of maximizing the expected profit and minimizing the risk of profit loss under worse conditions. However, these studies only deals with two-stage decision making, which is not suitable for long-term planing. This studies provides a framework for three-stage decision making. This model can use the mid- and long-term climate prediction, allowing the now-and-wait tradeoff and decision adjustment during the mid stage.

In scenario based stochastic programming, with a finite number of scenarios stochastic programming can be converted into an equivalent deterministic problem. Decisions should be based on data available at the time when the decisions are made and no future information

is needed. The properties of two-stage stochastic program are [22]: 1) decisions at all stage are made at once and no further information is expected, 2) hedging against all considered unrelated scenarios of possible developments is assumed, 3) except for the first stage, no non-anticipatively constraints appear. Depending on the considered problem in this study, such properties can be regarded as disadvantages. In this study, we extend the two-stage stochastic programming to a three-stage stochastic programming by permitting modified decisions in each stage based on the real-time realizations of uncertain system conditions [67, 1]. The uncertainties in stochastic programming are modeled as a 3-layer scenario tree described in section 4.1.

In our three-stage model, the "here and now" decision is the same in two-stage model, while the "wait and see" decisions are further extended by the short-term forecasting in 2040s and long-term forecasting in 2090s. The idea behind the framework is that if infrastructure capacity is not sufficient (i.e., BMPs expansion are limited), even the best tactical mitigation measures will not prevent large drought damage when a serious drought occurs; whereas excess infrastructure capacity means extra cost but less costly tactical measures may be sufficient to prevent a certain level of drought damage even under a limited infrastructure capacity. As the investment on strategic measures is higher than the tactical measures, given the climate forecasting, how investment is allocated between infrastructure expansion (strategic measures) and facility operation (tactical measures) for the preparedness and mitigation of drought under an uncertain future? How does the temporal allocation of investment depend on climate change projection? That is, how infrastructure is first planned based on mid-term climate forecasting and then adjusted based on long-term climate forecasting? Further more, how uncertainties from different climate forecasting horizons are handled in the decision framework? Compared with two-stage model, the three-stage model can not only tell the resources allocation between strategic and tactical measures, but also tell the temporal allocation between different planning horizons.

4.2.1 Model Formulation

The purpose of the decision making framework is to provide a quantitative tool for analyzing the tradeoffs between strategic and tactical drought mitigation measures. Strategic measures made for drought preparedness and tactical measures for drought mitigation can help reduce the damage of drought on agriculture and environment. If infrastructure capacity is not sufficient (i.e., BMPs sizes are limited), even the best tactical mitigation measures will not prevent large drought damage when a serious drought occurs; whereas excess infrastructure capacity means extra cost but less costly tactical measures may be sufficient to prevent a certain level of drought damage even under a limited infrastructure capacity.

The three stage stochastic model has 2 kinds of decision variables in two different periods, in figure 16. The decisions are: 1) BMPs building made at current time (bmp_1), 2) irrigation operational cost during the second stage ($irr_{s,yr}$, where s stands from climate scenarios in second stage, and yr represents the years in second period) , 3) BMPs expansion at the end of the second stage ($bmp_{2,s}$), 4) irrigation operational cost during the third stage ($irr_{st,yr}$, where st stands from climate scenarios in third stage led by scenario in second stage according to scenario tree). The second stage and third stage is represented from 2040-2049 and 2091-2100 respectively. The periods between the two stage is not considered in this study due to the climate data availability. The upper and lower bound of the decisions are the same described in chapter 3.

The objectives of the optimization problem are: 1) capital investment, 2) crop yield, and 3) low flow. The capital investment includes infrastructure expansion cost and irrigation operational cost in each stage. The BMPs investment cost is calculated as a point in the time horizon and irrigation cost is obtained from a period of 10 years in each stage. Considering the time value of capital investment, all the cost is converted into current value to with constant discount rate. The formulation of capital investment objective is:

$$Minimize\ c(bmp_1) + E\left(\sum_{yr=2040}^{2049}\left\{c\left[irr_{s,yr}(bmp1)\right]/(1+r)^{yr-2015}\right\}\right) \tag{4.1}$$

$$+E\left[c(bmp_{2s})\right]/(1+r)^{2050-2015} + E\left(\sum_{yr=2091}^{2100}\left\{c\left[irr_{st,yr}(bmp_1, bmp_{2s})\right]/(1+r)^{yr-2015}\right\}\right)$$

where r is the annual discount rate. $c(bmp_1)$ is the current (first stage) BMPs investment, which consists of infiltration pond and terrace expansion. Here, we made an assumption of equal probability distribution of all climate scenarios in the same stage. Thus, the expect value is taken according to the scenarios tree in section 4.1. There are some cost-efficient analysis on BMPs at different scales [26, 34, 3], but the cost of BMPs is different and highly uncertainty from case to cases. Arabi et al. [3] contribute BMP cost to establishment, maintenance, and opportunity costs. Establishment costs included the cost of BMP installation, and technical and field assistance. Maintenance cost is usually evaluated as a percentage of establishment cost. The opportunity cost is a dollar value that would be produced over the BMP design life as a result of investing the establishment and maintenance costs by purchasing saving bonds. In this study, we only consider the establishment cost of BMPs for simplicity. According to Kaini et al. [34], the cost is set at 500 per acre-ft and 300 per unit for infiltration pond and terrace respectively. The cost is only dependent on the size of BMPs, thus we don't consider the fixed cost such as transaction cost of BMPs. The BMPs expasion at the end of second stage is converted to current value by $(1+r)^{2050-2015}$.

The irrigation cost is the total irrigation cost over a 10-year length period during 2040s and 2090s. The irrigation cost depends on the groundwater table, total water pumped, labor cost and so on. In this study, we differentiate aquifer-based irrigation by different sources: shallow aquifer and deep aquifer. The irrigation cost is set at 10 and 20 per acre-inch for irrigation from shallow and deep aquifer, respective. The irrigation cost is only dependent on the amount of water pumped, so we ignore the fixed cost associated with the each irrigation events.

Crop yield can be converted into economic value. However, due to the uncertainty of future crop price, we treat it as a separate objective. This objective can be integrated into the first objective with reasonable crop price model, which is beyond the scope of this study. The second objective is to maximize crop yield during 2040s and 2090s, formulated as:

$$Maximize\ E\left[yld_2(bmp_1, irr_s)\right] + E\left[yld_3(bmp_1, bmp_{2s}, irr_{st})\right] \tag{4.2}$$

note $yld_2(bmp_1, irr_s)$ means the crop yield predicted by SVR during second stage with the expansion of BMPs in first stage and irrigation in second stage. The third stage crop yield is affected by BMPs from first and second stage, and irrigation in third stage.

The third objective is for environmental conservation for in-stream ecology. We choose to maximize 10-day-minimal flow over the planning periods:

$$Maximize\ E\left[flow_2(bmp_1, irr_s)\right] + E\left[flow_3(bmp_1, bmp_{2s}, irr_{st})\right] \tag{4.3}$$

For the latter two objectives, we do not discount the future value, meaning we put equal weight on agricultural production and environmental conservation from different periods.

The decision-making framework is a coupled simulation-optimization model. The constraints for the optimization model is that all the state variable should based on the prediction of SVR model described in chapter 3:

$$SVR_{2s}\left[bmp_1, irr_s, yld_2(bmp_1, irr_s), flow_2(bmp_1, irr_s)\right] = 0 \tag{4.4}$$

and

$$SVR_{3t}\left[bmp_1, bmp_{2s}, irr_{st},\ yld_3(bmp_1, bmp_{2s}, irr_{st}),\ flow_3(bmp_1, bmp_{2s}, irr_{st})\right] = 0 \quad (4.5)$$

Note the superscripts s and t represent the relation of scenarios in second and third stage. The links are illustrated in figure 16. For example, only $t = 1, 2 and 3$ scenarios in third stage can be led by $s = 1$ scenario in second stage. The scenario tree are represented by the superscripts:

$$t = \begin{cases} 1, 2, 3, & when\ s = 1 \\ 4, 5, & when\ s = 2 \\ 6, 7, & when\ s = 3 \\ 8, 9, & when\ s = 4 \end{cases} \quad (4.6)$$

4.3 Solution Algorithm

As SVR coupled into the optimization problem, there is no explicit mathematical equation expression for the relations. We use evolution algorithm to solve the optimization problem. The problem is encoded into MATLAB ® and solved by multi-objective genetic algorithm(multi-ga) in optimization toolbox. The population size is set at 200 and generation is set at 150. The crossover probability is set at 0.85. Other multi-ga parameters are set at their default value.

4.4 Optimization Results

As the third stage is about 80 years from now, the discount rate has a large effect on the optimal solution, we solve the problem with different discount rate from 1% to 9%. Then further analysis is discussed when discount rate is at 4%.

4.4.1 Effect of Discount Rate

In the investment objective 4.1, the future costs decay exponentially with the time from current. Large discount rate converts future value into small current cost. The multi-objective GA produces a Pareto frontier for each optimization at the three objective dimensions. To analysis the effect of discount rate, we fix two objectives: low-flow at 0.6 m^3/s and yield at 8 ton/ha and pick out the current value cost in the Pareto frontier, shown in figure 18.

The current cost value decreases from 350 million $ to 90 million $ when discount rate increases from 1% to 9%. Note that the decreases is the current value, it does not necessary the BMPs expansion scale is smaller. Discount rate at 4% can be regarded as a turning point, below which the current cost value decreases significantly with discount rate and above which the current cost value is not so sensitive to discount rate. So 4% discount rate is chosen for further analysis.

Similar to figure 18, figure 19 and 20 is obtained with low flow and crop yield objectives fixed. Figure 19 shows with the same lowflow and crop yield requirement, how discount rate affects the BMPs investment between second and third stage. Here the BMPs cost is calculate at each stage and not converted into current value, thus shows the actual BMPs scale. The result shows with increasing discount rate, BMPs investment for first stage decreases and that for second stage increases. The investment on BMPs shifts to latter years as the current value of second stage BMPs investment significantly decreases with large discount rate. In figure 20, irrigation increases with discount rate because the higher discount rate means less expensive future irrigation. The investment on BMPs in second stage is about 2 to 4 times of that in first stage; irrigation cost during 2040s is about 2 time of that in 2090s. We can see both the BMPs expansion and irrigation operation for 2090s is higher than these from 2040s. More investment should be spent for mitigation the drought in 2090s, as the climate model predicts drier conditions in 2090s.

4.4.2 Optimization Results with 4% Discount Rate

In figure 18, we have seen that 4% discount rate shows a changing behavior of cost in current value on discount rate. Below this discount rate current cost value decreases significantly with discount rate and above this discount rate the current cost value is not so sensitive to discount rate. We pick up the optimization results with 4% discount rate for further analysis. Figure 21 shows the Pareto frontier of three objectives, which compromises a surface in the 3 objectives dimension.

For more detailed analysis, we choose the points in Pareto frontier with different current cost value. As the Pareto frontier does not provide the exact current cost value we choose, we pick up the 4 closest points to our specified current cost value (60,70,80 and 90 million $, respectively). This is to use a plane parallel to 'yield'-'lowflow' plane to cue the Pareto frontier, and extract the intersection points. As there is no exact intersect between points and plane, the points close to the intersection plane is projected on the intersection plane from Z axle. In figure 22, each line is the a Pareto font between yield and low-flow objectives. For example, then current cost is 90 million $, trade-offs are made between yield and low-flow along the line. With fixed budget, decision makers can invest more on BMPs to have a higher low-flow, which leads to less investment on irrigation investment; or decision makers can put more on agricultural production at the expense of low-flow. This provide the information for decision makers to compromise between crop profit and environmental conservation. Although we only pick 4 points, we can see the the slope of the trade-off line changes. The tradeoff line have steep slope at high yield (or small low-flow) part and becomes flat at at low yield (or large low-flow) part. The slope has a big change around 0.2 m^3/S low-flow rate. At high crop yield section (e.g. 8 ton/ha), the increased investment does not help to increases low-flow much. If we reduce the yield a little, there is a significant environmental benefit. For example, if the yield crop is reduced at 6.8 ton/ha (red line in

figure 22), low-flow rate increases from 0.17 to 0.7 m^3/s with investment increasing from 60-90 million dollar, while the increases is only from 0.03 to 0.2 m^3/s at high crop yield (8 ton/ha). The tradeoff line tells us when investment for in-stream flow sustainment is effective according to the crop yield. We can also see with increasing budget, the trade-offs line moves up-right ward, which means more capital investment produces results with higher yield and higher low-flow at the same time.

To see the how investment is split for BMPs and irrigation at different stage, we calculate the mean BMPs investment cost and irrigation operation cost (these value are not converted into current value) from the trade-offs line, shown in figure 19 and 20. The future investment value represents the actual size of these BMPs. The BMPs implemented in 2050 are about 4 times larger than these planned in first stage. This means instead of putting all the investment at now, the optimal solution indicates we should wait until the 2050 and have a large scale BMPs expansion to prepare for the drought in 2090s. Although BMPs can mitigate the damage of drought, excess infrastructure capacity is not necessary in 2040s because the less costly tactical measures are already sufficient to prevent a certain level of drought predicted by the climate model. However, even the best tactical mitigation measures are not enough to prevent the damage by the serious drought during 2090s, and large infrastructure capacity needs to be extended in advance for preparedness. The irrigation amount in 2090s are about 1.5 time of that in 2040s. Although droughts are more frequent in 2090s, large scale infrastructure expansion helps to reduce the dependence of irrigation to mitigate the drought. With a increasing budget, the investment is shifting from tactical measures to strategic measures. We can see that how capital investment is split between infrastructure expansion and irrigation operation with different investment rate. The irrigation cost does not change quite much with increasing investment in figure 20, while investment on BMPs increases with capital investment. The BMPs investment increases from 28 to 40, and 140 to 176 for first stage and second stage respectively, in figure 19.

To see how should decision makers invest infrastructure now or wait-and-see for each measures, given the uncertainties from mid- and longterm climate change prediction, we compare the cost in current value of investment in different stages, in figure 25. Note the costs are all converted into current value, this represents how to allocate investment, not the physical scale of different measure. The investment for infrastructure takes more than half of the total investment. As the budget increases, the investment on irrigation decreases (especially irrigation for 2040s) and more investments goes to BMPs. BMPs built in 2050 takes the largest portion, taking 34% to 40% of the total investment when the budget increases from 60 to 90 million dollar. The investment on BMPs now and future in total takes 60% to 78% of the whole investment, figure 26.

4.5 Summary

This chapter provides a decision framework for investment allocation between strategic measures (e.g.,infrastructure expansion) and tactical measures (e.g., facility operation). The three-stage stochastic optimization shows how decision makers would invest infrastructure now or "wait and see", given the given the uncertainties from mid- and long term climate change prediction. The uncertainties from mid- and long term climate change prediction are handled in a scenarios tree, in which climate predictions in different stages are formed in a tree structure based on the assumptions in CO_2 concentration levels. This bundling by assumptions averages the uncertainties from climate sensitivity of RCM-GCM models.

The optimization results show that although BMPs can mitigate the damage of drought, excess infrastructure capacity is not necessary in current decision as the less costly tactical measures are already sufficient to prevent a certain level of drought during 2040s predicted by the climate model. However, even the best tactical mitigation measures are not enough to prevent the damage by the serious drought during 2090s, and large infrastructure capacity

needs to be extended in advance for preparedness. Thus larger BMPs expansion is proposed in 2040s for droughts preparedness in 2090s.

Chapter 5

Conclusions

5.1 Discussion on Results

This study provides a decision making framework of infrastructure expansion and facility operation for drought mitigation under climate change. Mid- and long-term climate forecast are provided by regional climate model dynamically downscaled from global climate model. The predicted future annual precipitation ranges from 305.5 mm to 594.6 mm, compared with 519.4 mm in current value. 2 out of 12 scenarios predict increasing annual precipitation and 2 out of 12 predict unchanged annual precipitation, while others predict decreasing annual precipitation. Hadley model predicts precipitation increase in spring and decrease in other seasons. CCSM model predicts decreasing precipitation in all months except June. The decreasing growing season precipitation indicates more irrigation need for agriculture production. For the temperature, all model predict a increase, ranging from 0.6 to 9.2 ^{o}C. The high temperature increases the crop evapotranspiration, and the heat wave caused by extreme high temperature(e.g. in CCSM predictions) also reduces crop yield even if there is enough water.

A quasi-physically based hydrological agronomic watershed model is applied in FCB. The model is built on SWAT with modifications. Although the original SWAT has groundwater based irrigation, the baseflow is calculated separately from the aquifer storage, thus not affected by pumping. We replace the baseflow calculation with a conceptual linear groundwater reservoir model and get baseflow proportionally from aquifer storage. By updating aquifer storage (including pumping) every time step, the modified model can simulate the

interaction between pumping, aquifer storage and baseflow. The model is calibrated and validated with historical data and then used to predict the watershed response to infrastructure expansion and facility operation under different climate scenarios. Due to the computation expense, a statistical surrogate model is built with SVR to replace SWAT in the optimization. Each SVR is trained from the inputs (parameters representing infrastructure expansion and facility operation) and outputs (irrigation amount, low-flow and yield) from SWAT model with each climate scenarios. The SVR simulation predicts that infiltration pond and terrace are effective BMPs to increases the low-flow during drought conditions. However, there are different patterns in infiltration pond v.s. low-flow relationship. For slightly dry scenarios, low-flow monotonically increases with pond size; for extremes dry scenarios, the relation shows a convex pattern (increases first and then decreases). The nonlinearity is caused by the complicating effects of groundwater recharge and pond evaporation. During extreme drought events, large ponds evaporate water rather than recharging into aquifer. Irrigation eases crop water stress, while crop potential yield is reduces by heat wave in extreme dry scenarios even there is enough water. Irrigation from shallow aquifer causes stream depletion, but can be recovered by recharges from BMPs. Irrigation from deep aquifer increases low-flow as deep aquifer is hydraulically unconnected with stream, but is more costly and requires long time to recover. Shallow aquifer irrigation is enough for moderate drought mitigation, and deep aquifer irrigation is essential for crop production and low-flow preservation during severe drought.

For a watershed scale drought preparedness and mitigation, a scenario based three-stage stochastic optimization model is built to analyzes the relative role of strategic measures (e.g., infrastructure expansion) compared to tactical measures (e.g., facility operation) for drought preparedness and mitigation under climate change. To reduce the uncertainty of different RCM, the mid-term (2040s) and long-term (2090s) climate forecasting scenarios are bundled by their assumptions on CO_2 concentration level and further developed into a scenario tree. This framework gives information of here-and-now or wait-and-see decisions

on infrastructure expansion given the uncertainties from mid- and long-term climate change prediction. The model is to maximize crop yield and low-flow given the capital investment. The model shows that 1) even the best tactical measures (irrigation operation) alone are not sufficient for drought mitigation in the future and infrastructure expansion is critical, especially for environmental conversation purposes. With increasing capital budget, investment shifts from tactical measures (for drought mitigation) to strategic measures (for drought preparedness), 2) for a fixed capital investment, trade-offs should be made between crop production and environmental conservation. A small reduce in crop yield has significant benefit on low-flow preservation, while increases in low-flow is insensitive when crop yield is very low, 3) For the temporal allocation of investment, infrastructure expansion is preferred for the long term plan than the mid-term plan. Larger investment is proposed in 2040s than the current, due to a larger likelihood of drought in 2090s than 2040s. Moderate BMPs expansion is proposed now to prepare for drought in 2040s, and after the 2040s climate unfolds, large BMPs expansion is proposed for droughts preparedness in 2090s.

5.2 Further Work

In this study, the future climate is represented in mid- and long-term forecasting from RCMs, each with a 10-year length. This time-slice approach fails to account for the transient nature of climate changes over century. Climate forecasting within each period are stationary rather than transient, thus uncertainties linked to natural climate variability if not explicitly considered. Moreover, due to the data availability, we use 10-year length time series to represent mid- and long-term climate forecasting rather than a continuous climate forecasting for the whole time domain. However, groundwater related components have long time response. Thus, we do not consider the cumulative effect of infiltration pond recharge and groundwater pumping. There are two ways to address the problem. One is very straightforward: using continuous climate forecasting if there are available RCM results. However, continuous

RCM simulations are based on CO_2 concentration level at the end of the century. This approach only solves the continuous simulation of groundwater, while the mid- and long-term forecasting are based on the same CO_2 concentration level assumption. Another approach is to generate transient climate data from weather generator. Recently, Rascal et al. develops a transient weather generator, which enables the stochastic generation of large numbers of equiprobable climatic time series, representing transient climate change between two periods [27, 11]. This approach provides a continuous transient climate change over the whole century and also enables the climate change impact on groundwater resources. However, provided the continuous climate change data, the simulation model computation expense would be extremely huge. Now current computation technique, such as cloud computation, may help to solve this problem.

The scenarios based stochastic optimization in this study minimize the expected value of a cost function or maximize a net benefit function, and do not facilitate evaluation of the trade-offs between the risks of infeasibility and the losses in optimality. In addition, we assume a equal probability of the scenarios. Although there are other rules to determine the uncertainty of climate models such as giving higher priority to models more strongly verified by the historical observation or A maximum entropy method [36], the probability itself should be recognized as uncertain. Robust optimization (RO) can be used to incorporate risk aversion into optimization models. For a scenarios-based RO, higher moments (e.g., variance) are then introduced into the objective function as a measure of risk [46]. This allows the evaluation of trade-offs between the expected value of the objective function, the variability in the value of the objective function, and the risk of violating soft constraints [32, 12]. However, RO studies are based on single objective problems. Methods to handle the risk in a multi-objective problem need further study. Still, one of the disadvantages of RO is the potential size and complexity of the resulting model [66]. As a result, special solution algorithms may be required. Another promising way is the nonprobabilistic RO approach [6]. In this method, the uncertainty is not described by a PDF or scenarios but is known to

reside within a user-defined uncertainty set. Hence, instead of immunizing the solution in a probabilistic sense, the decision-maker searches for a solution that is optimal for all possible realizations of the uncertainty set. If there is a way to construct the uncertainty set from the climate forecasting, the decision framework avoids the assumptions on PDF of the scenarios.

Chapter 6

Figures

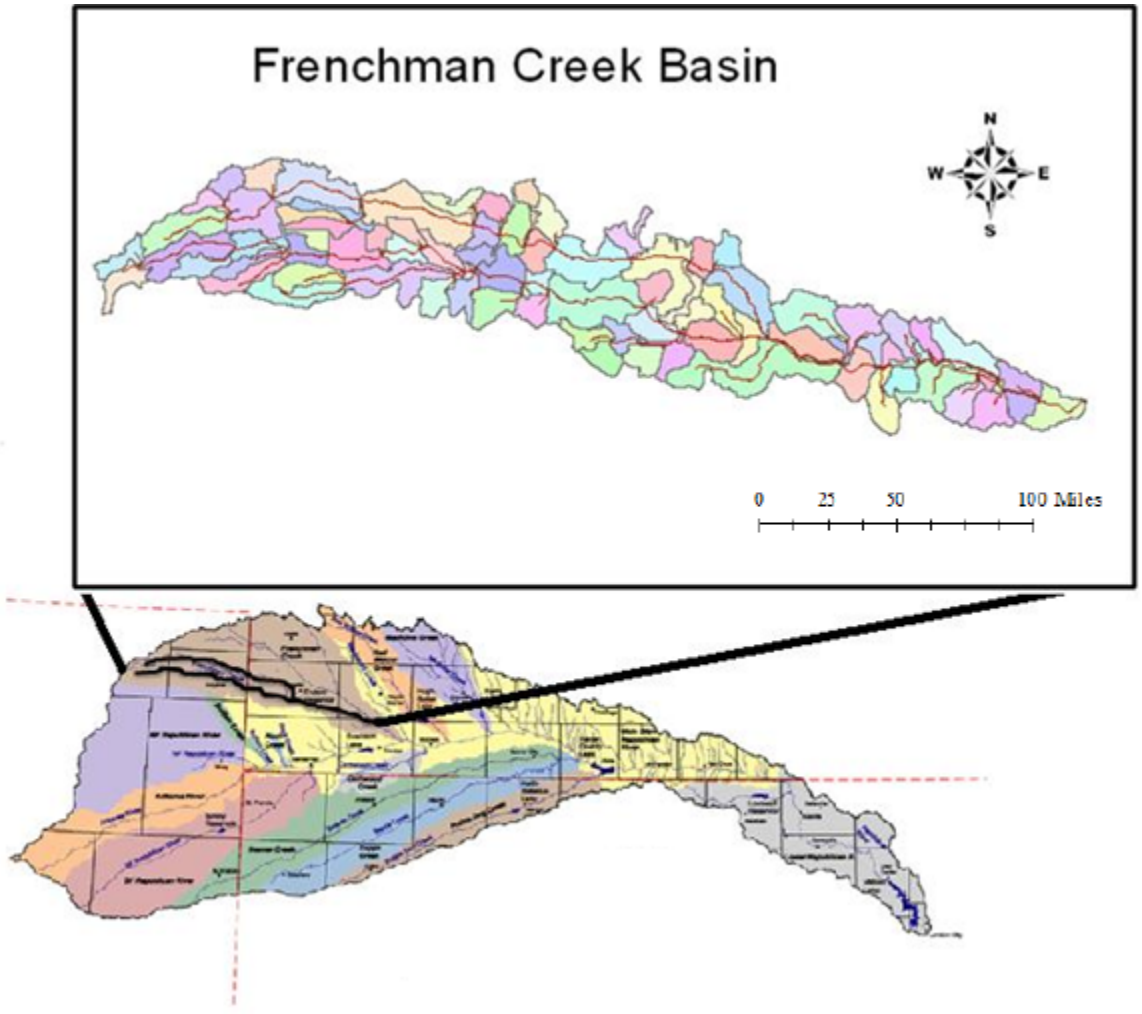

Figure 1: Domain of Frenchman Creek Basin in Republican River Basin

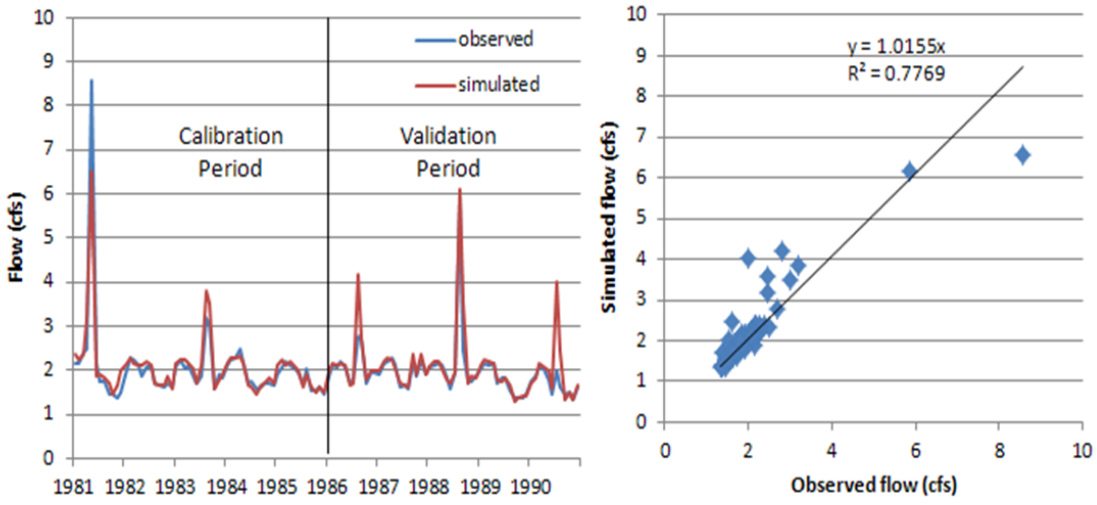

Figure 2: Calibration (1981-1985) and validation (1981-1985) of streamflow

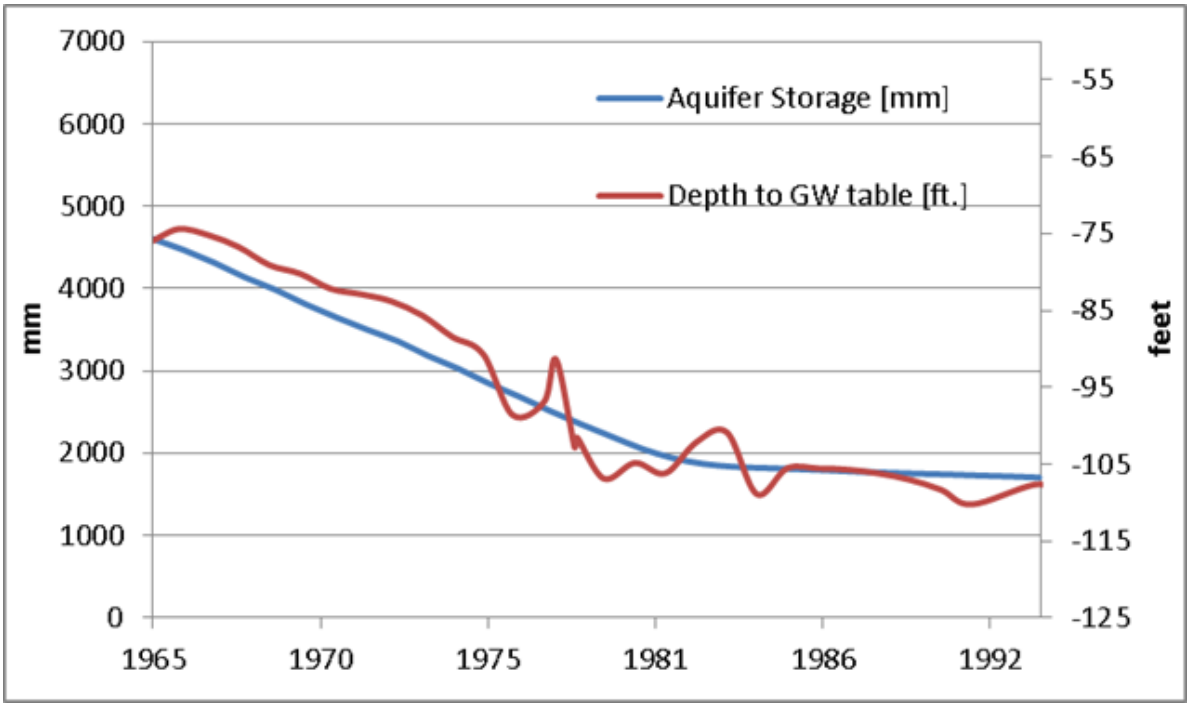

Figure 3: Comparison of simulated aquifer storage and USGS groundwater table height

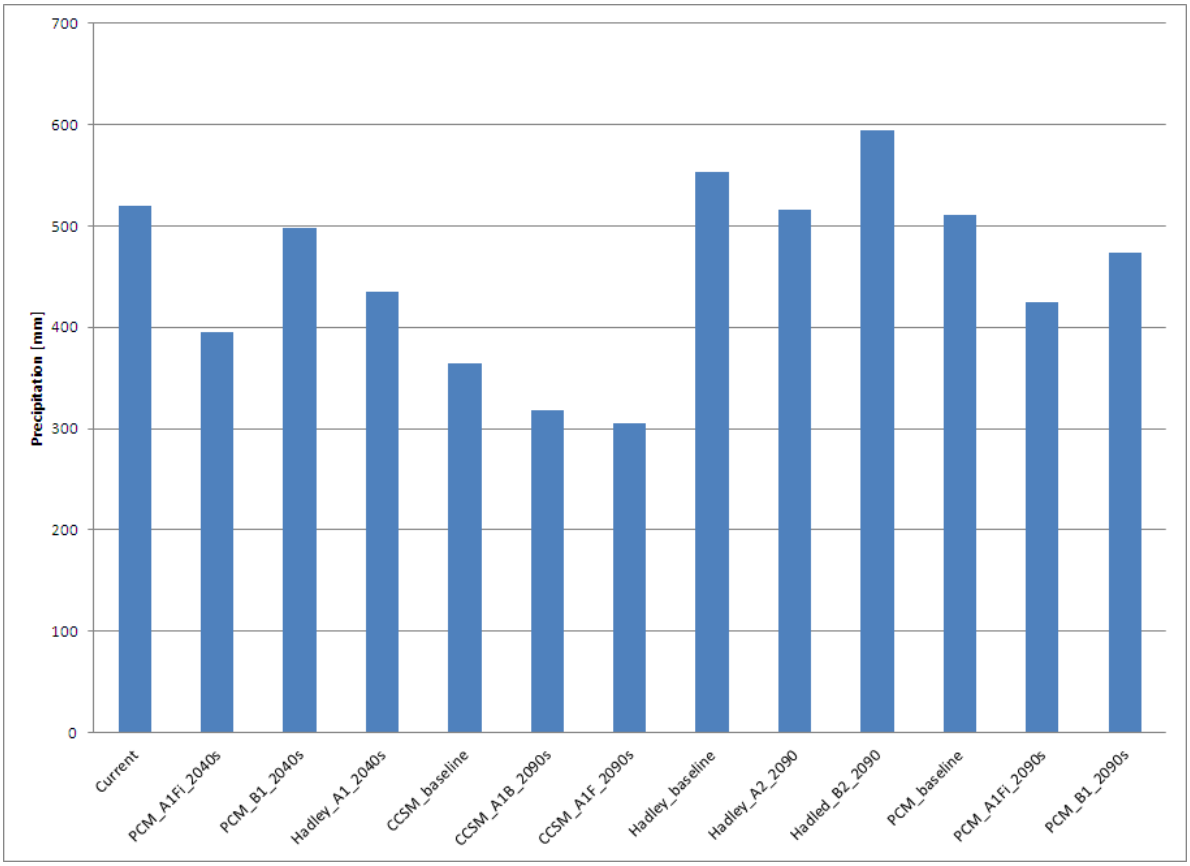

Figure 4: Annual precipitation predicted by different scenarios

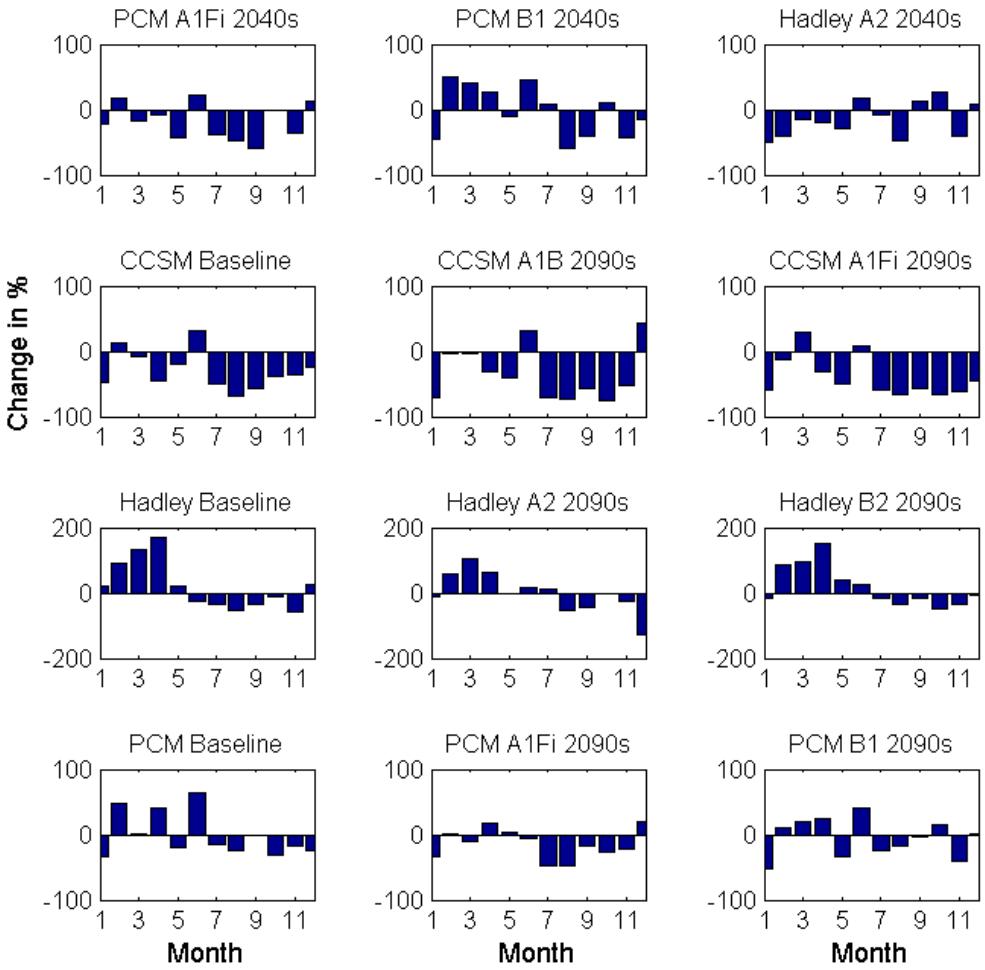

Figure 5: Monthly precipitation percentage change predicted by different scenarios

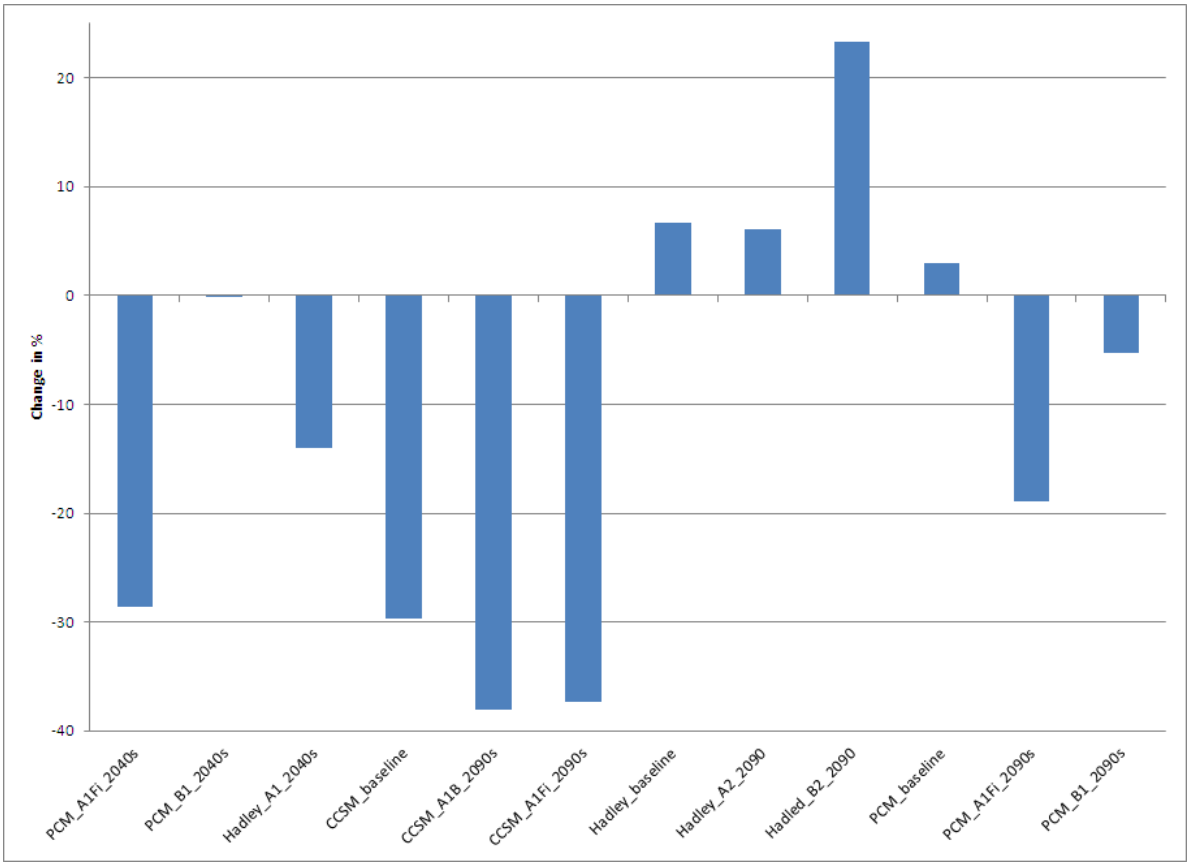

Figure 6: Growing season precipitation percentage change predicted by different scenarios

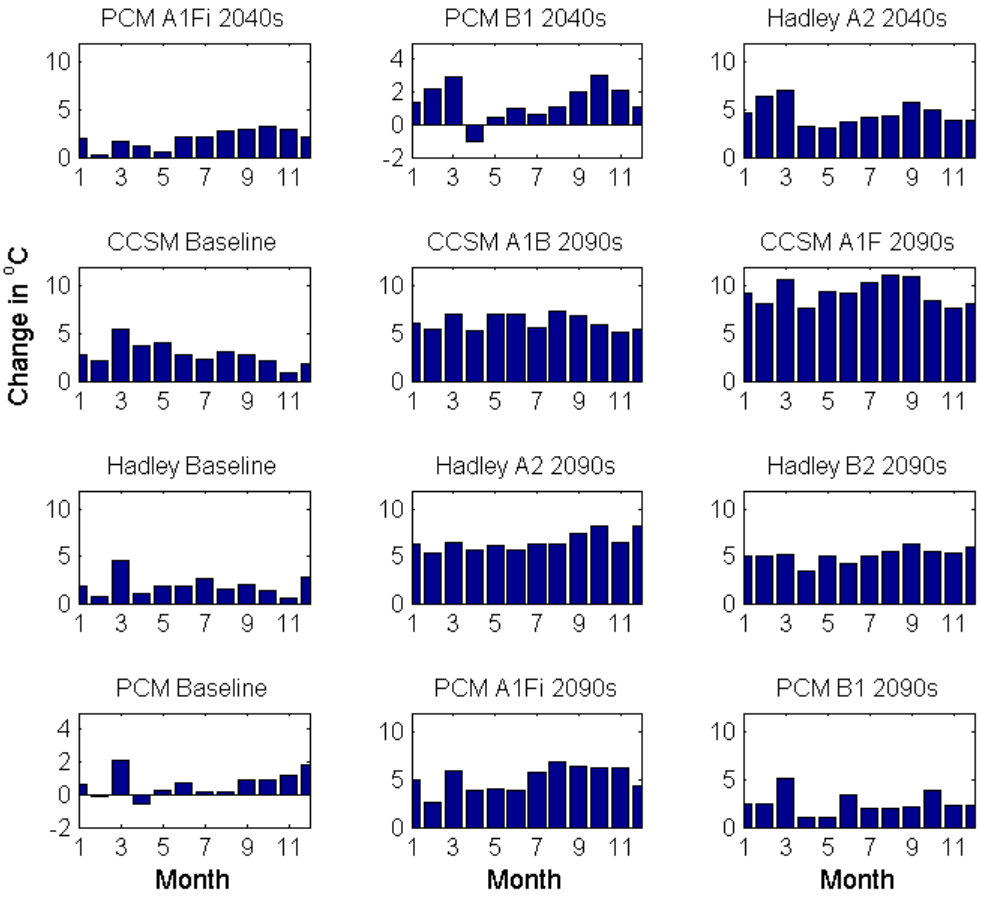

Figure 7: Monthly temperature change in °C predicted by different scenarios

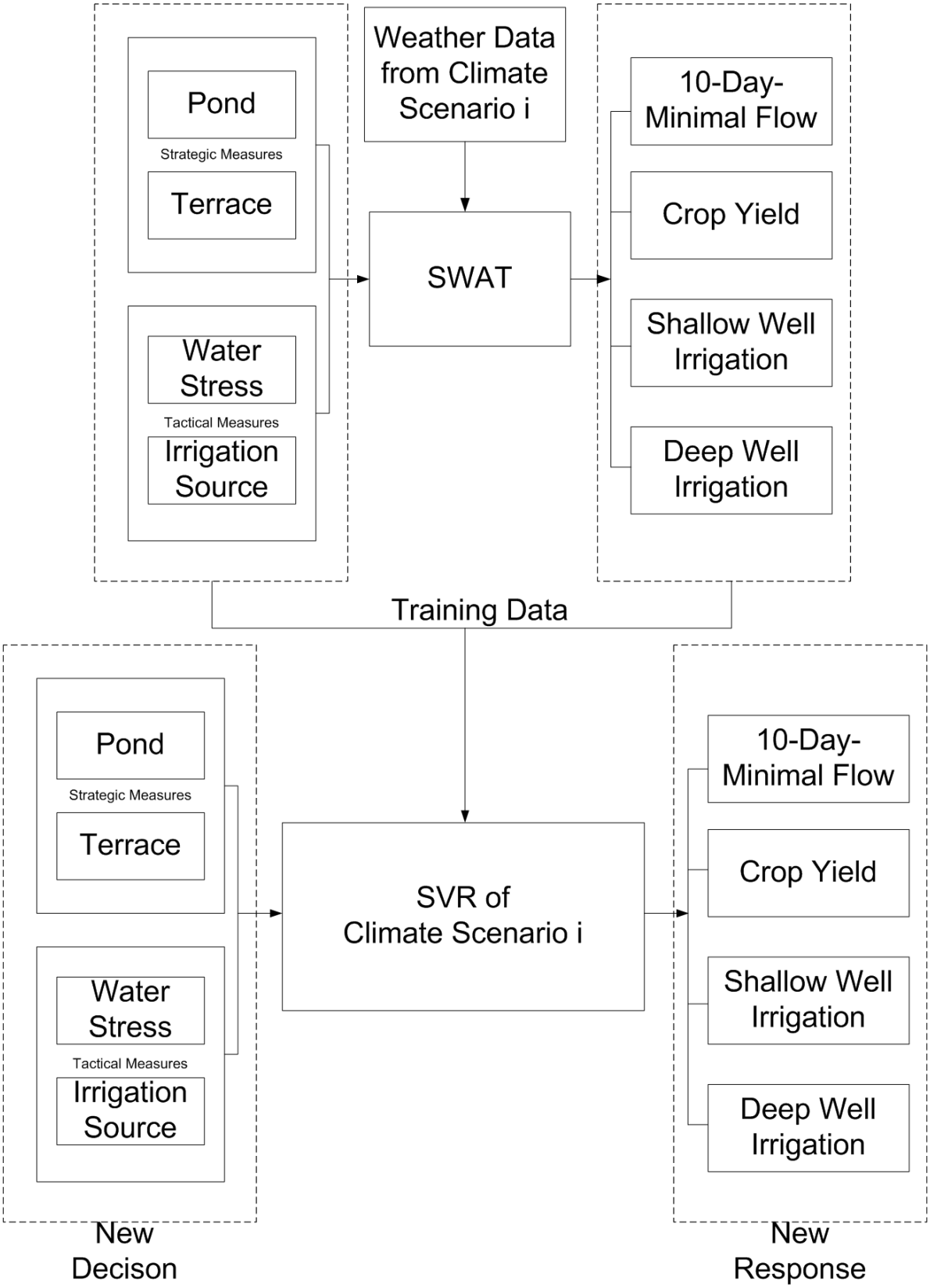

Figure 8: Processes of build SVR from SWAT model for each climate scenarios

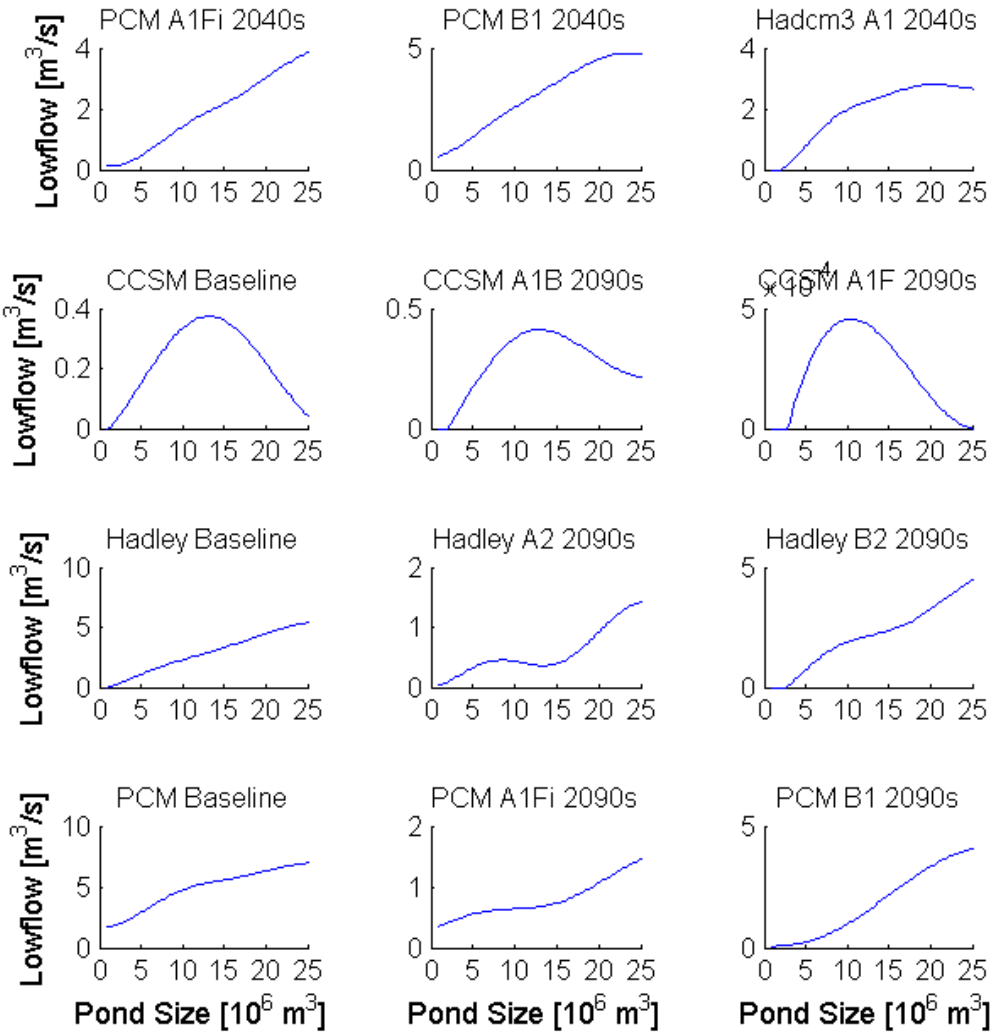

Figure 9: Effect of infiltration pond size on 10-day-flow under different climate scenarios

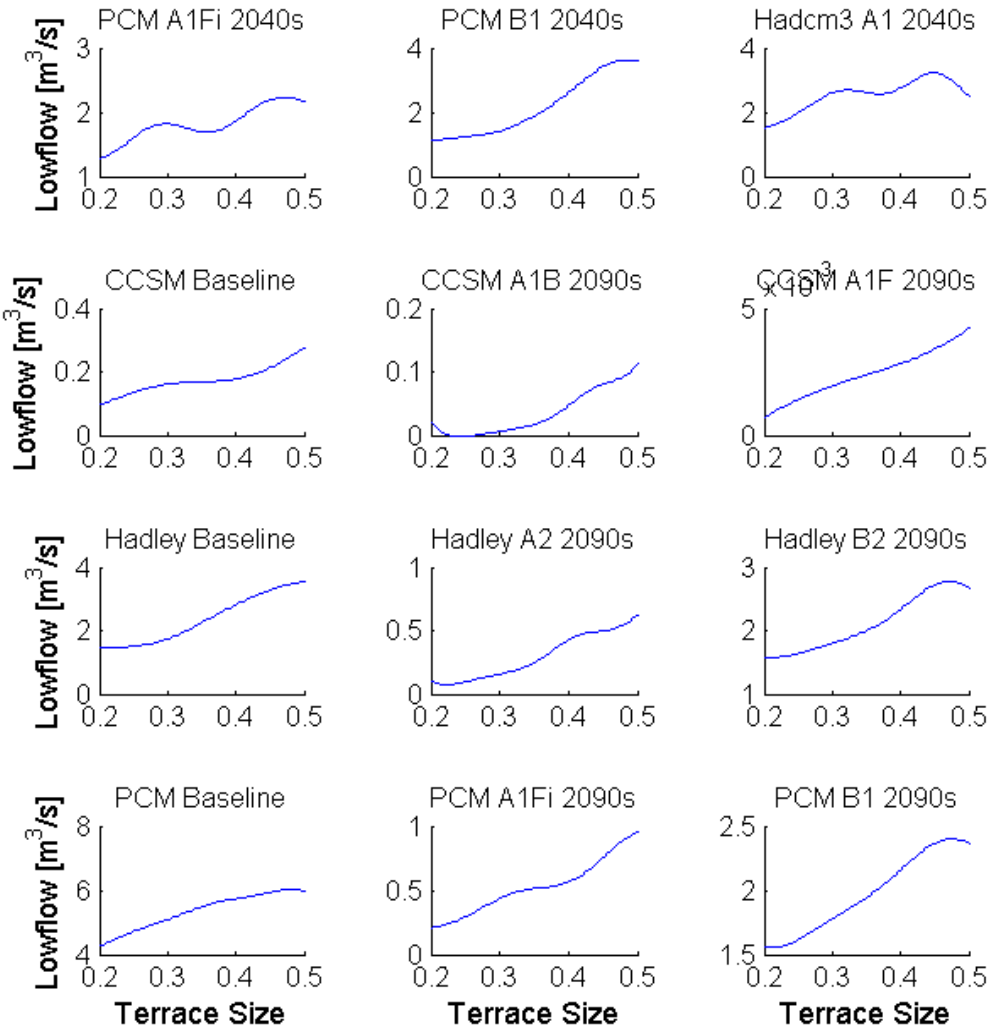

Figure 10: Effect of terrace expansion on 10-day lowflow under different climate scenarios

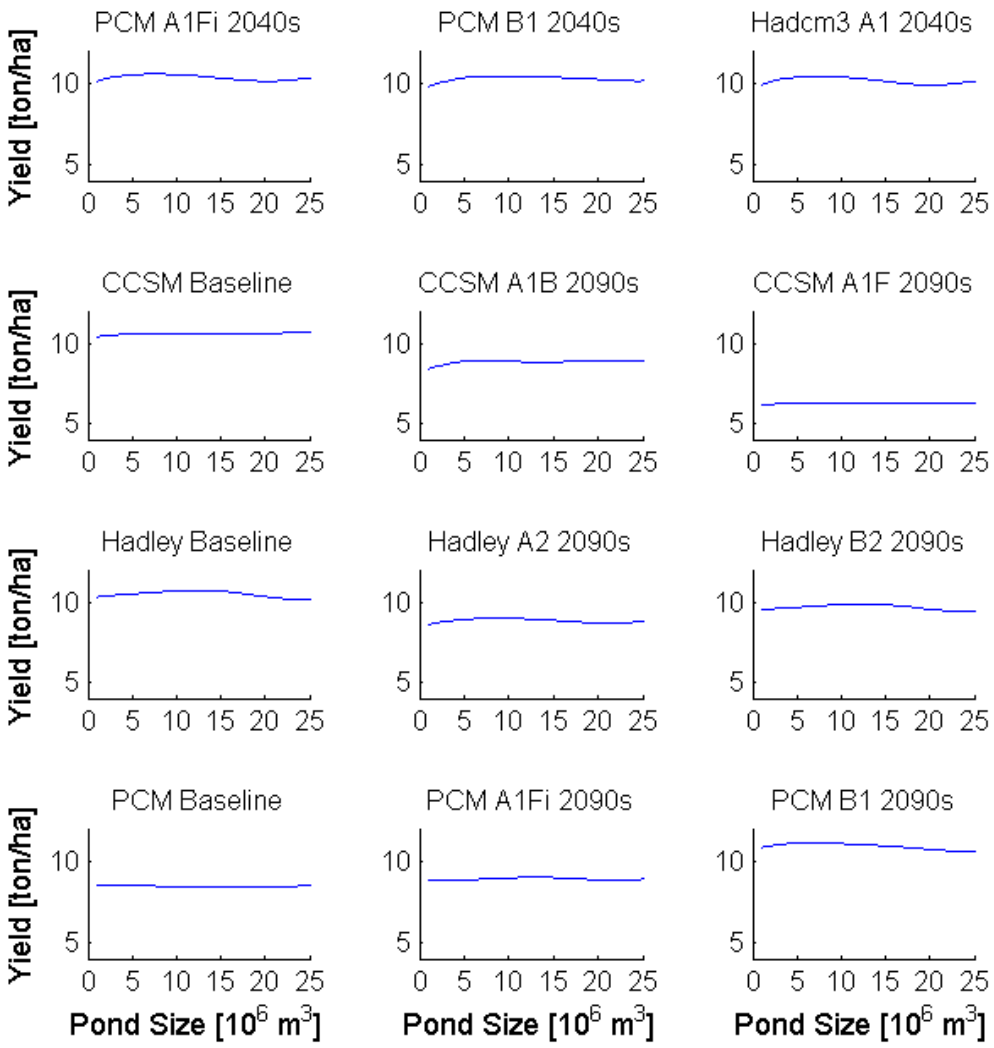

Figure 11: Effect of infiltration pond size on crop yield under different climate scenarios

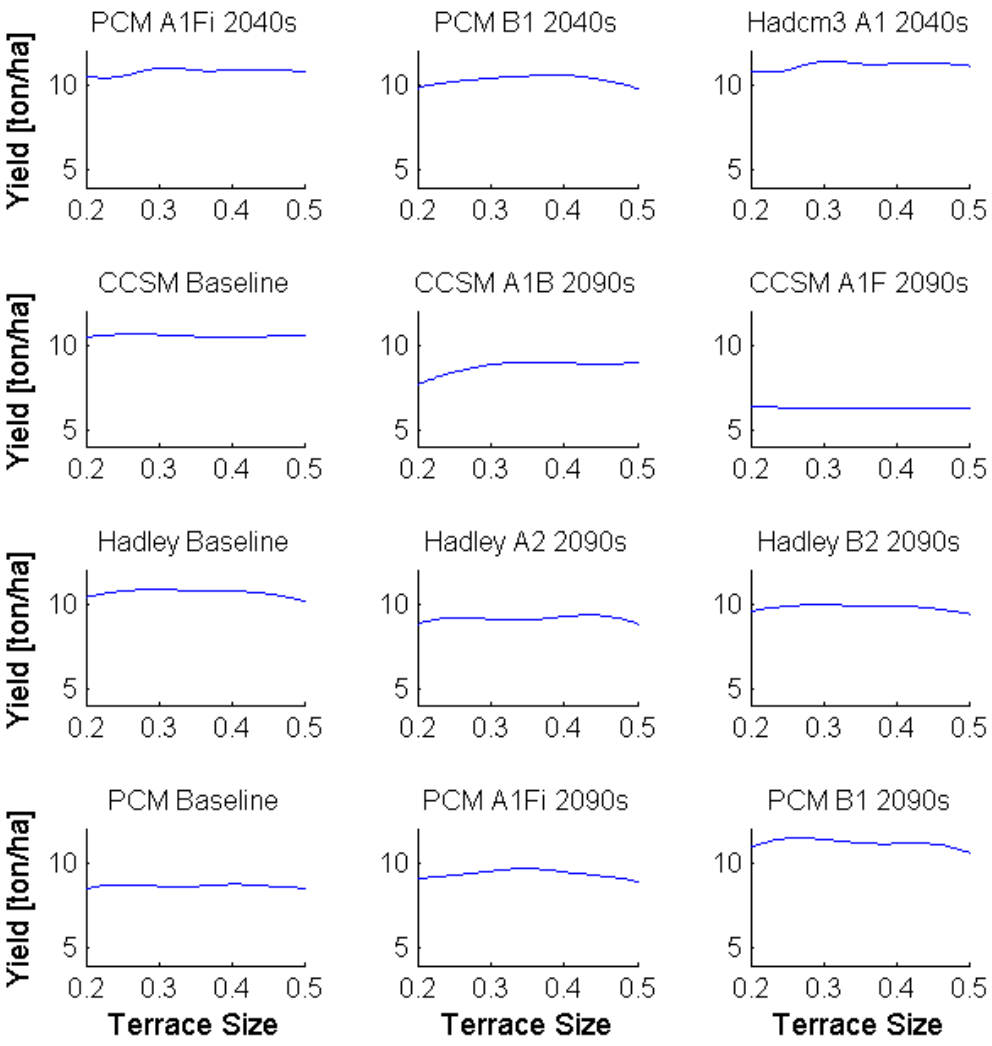

Figure 12: Effect of terrace expansion on crop yield under different climate scenarios

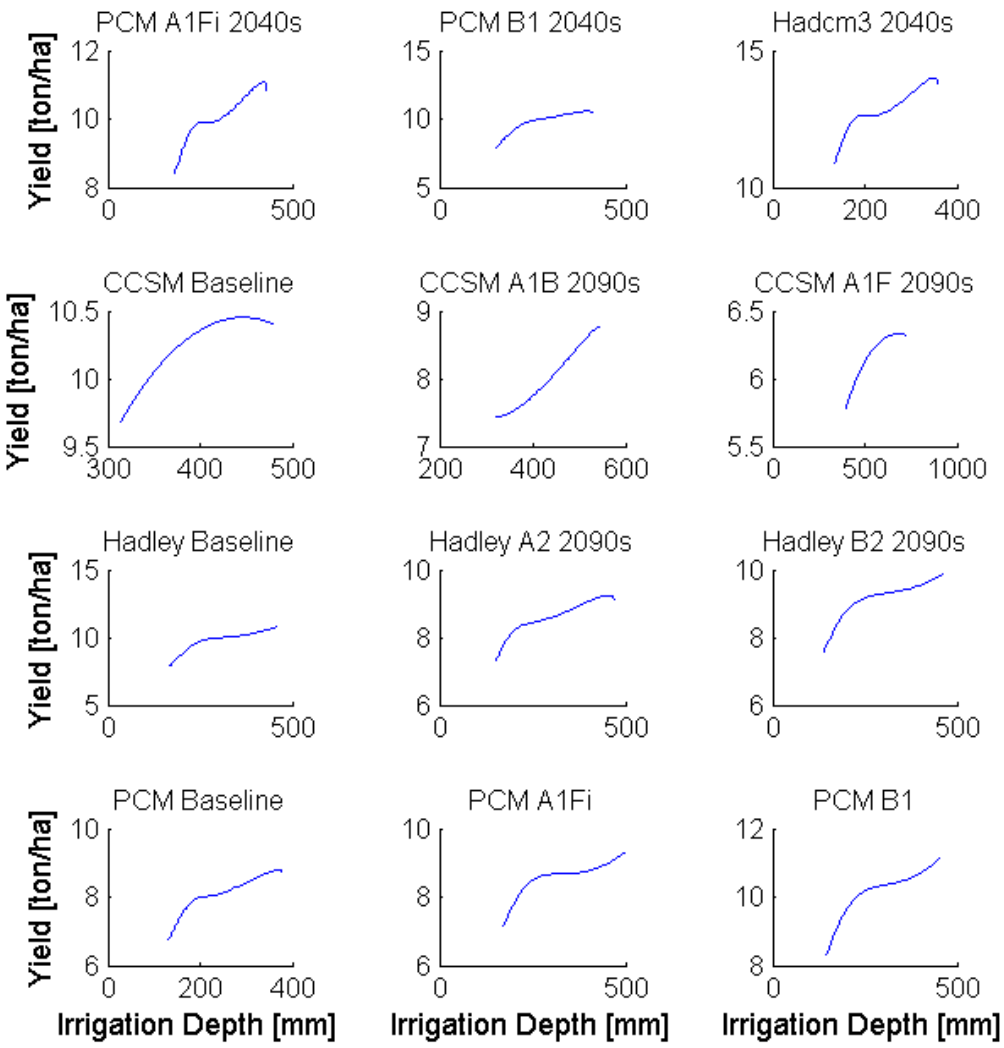

Figure 13: Effect of irrigation amount on crop yield under different climate scenarios

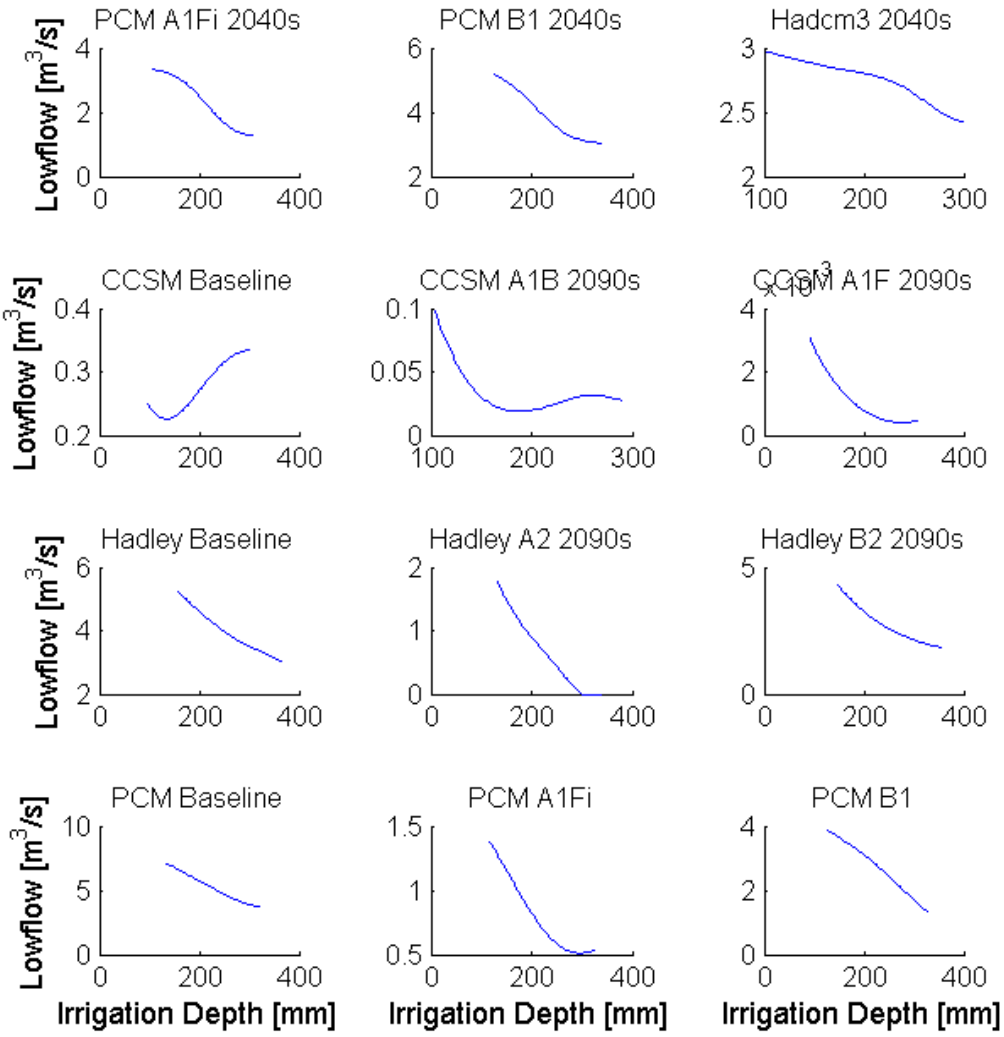

Figure 14: Effect of irrigation from shallow aquifer on 10-day-flow under different climate scenarios

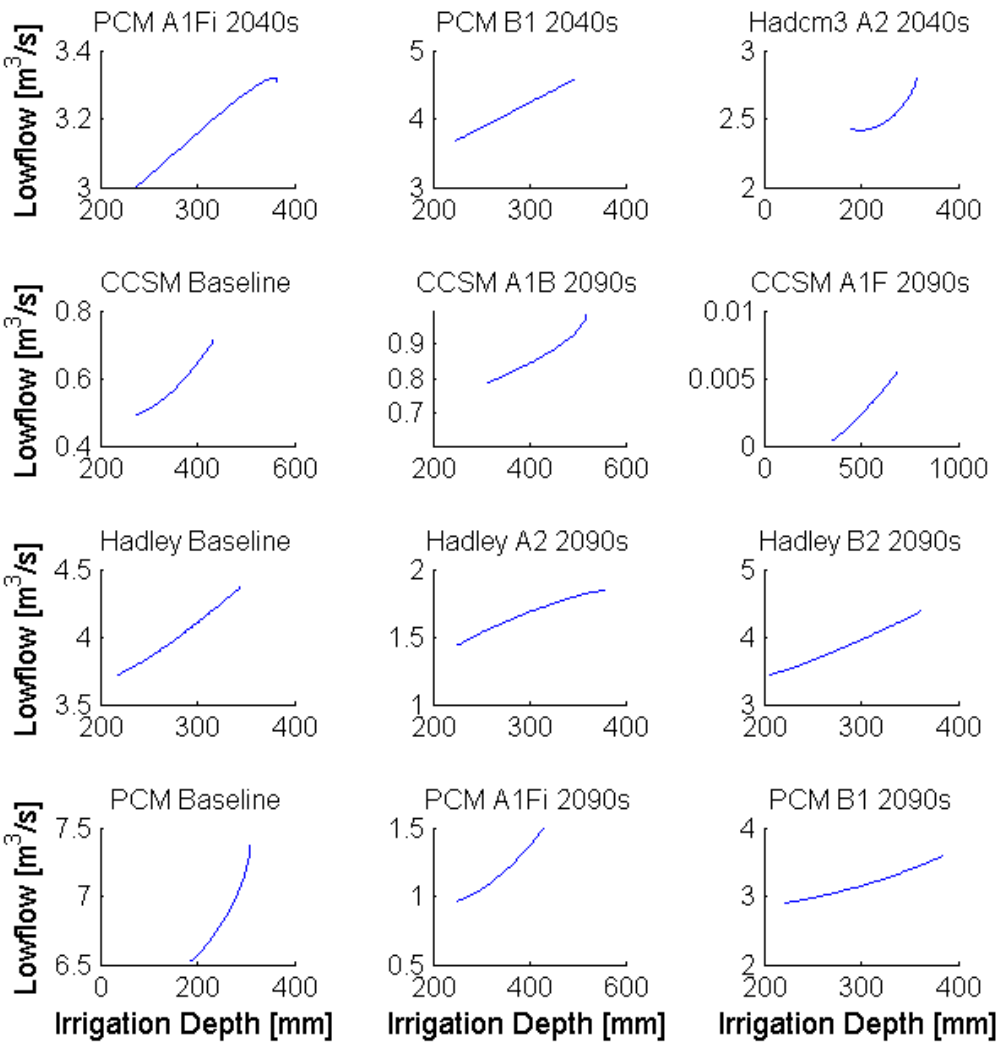

Figure 15: Effect of irrigation from lower aquifer on 10-day-flow under different climate scenarios

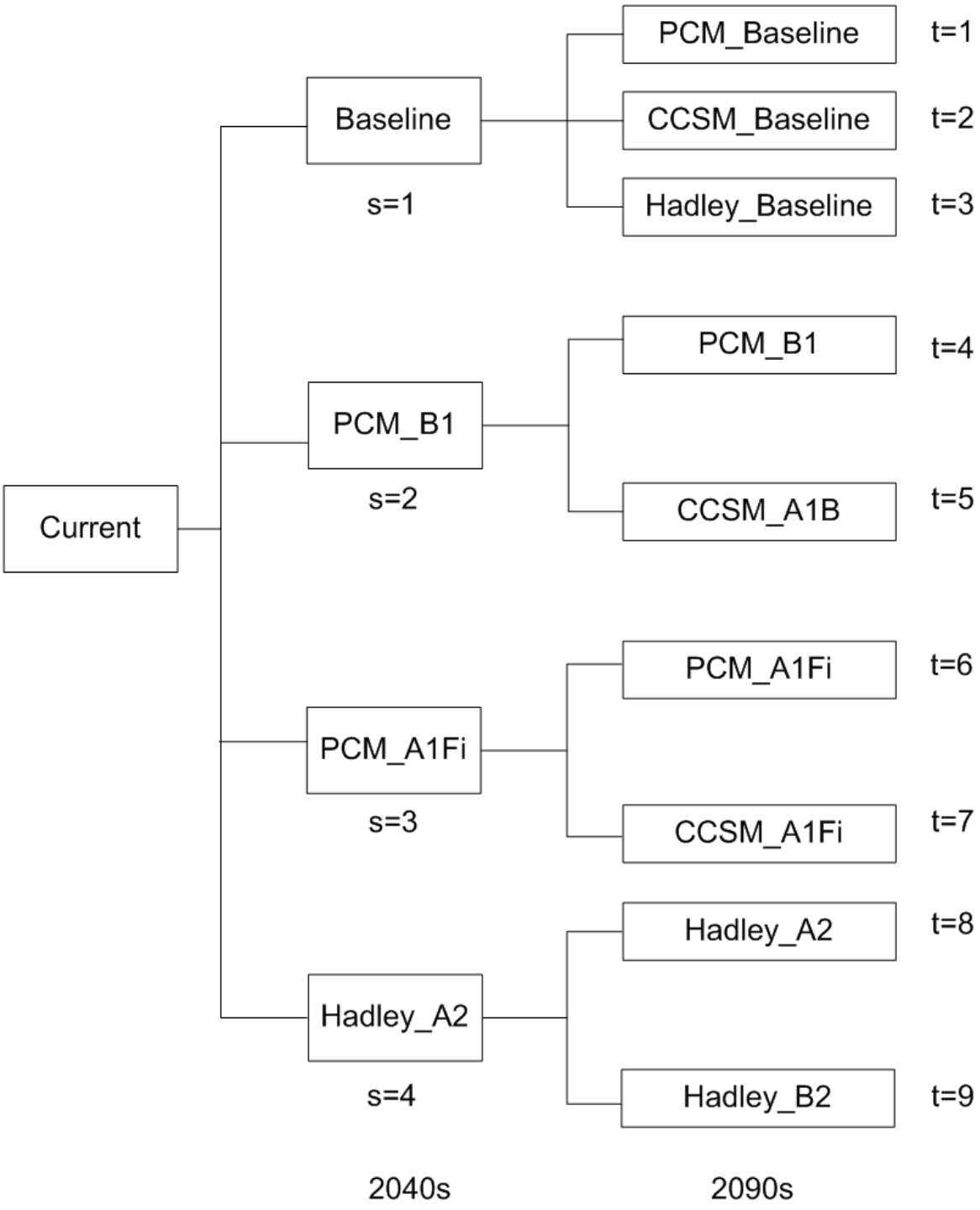

Figure 16: Scenario tree of three stage programming

.

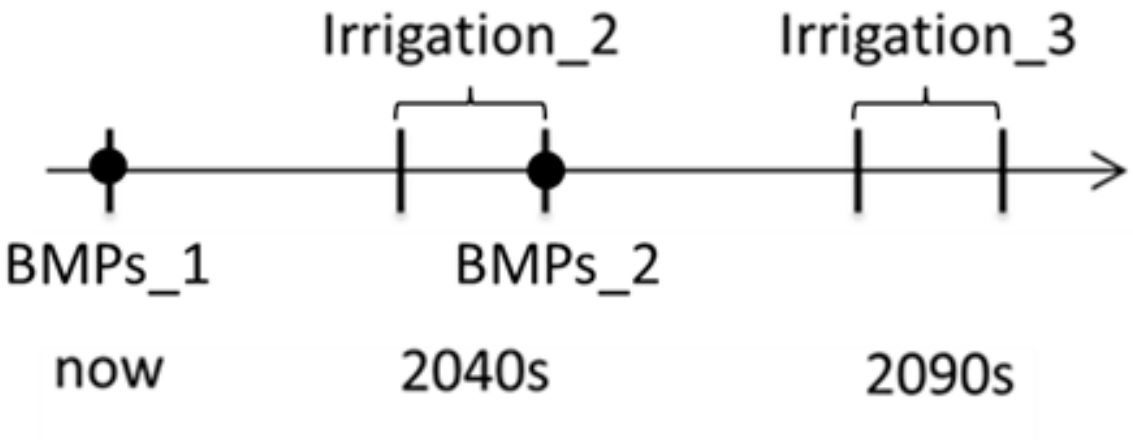

Figure 17: Three stage optimization model decision variables

.

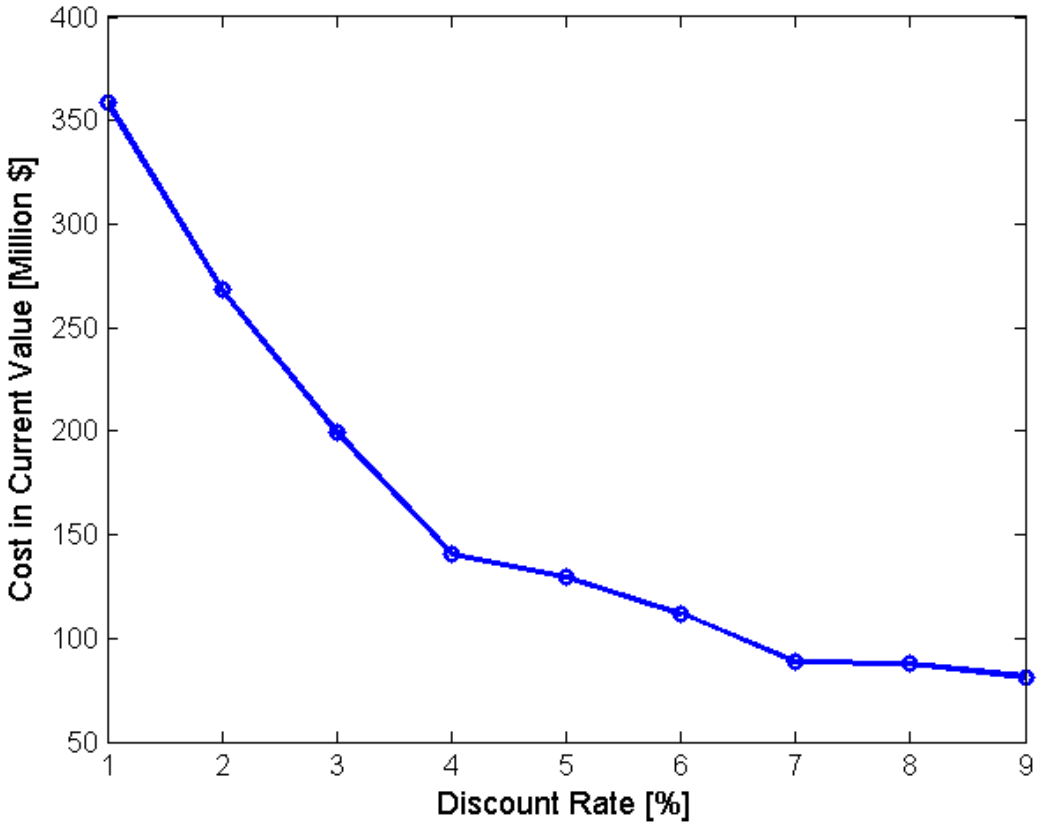

Figure 18: Cost in current value changes with discount rate when low-flow and crop yield objectives are fixed at 0.6 m^3/s and 8 ton/ha, respectively

.

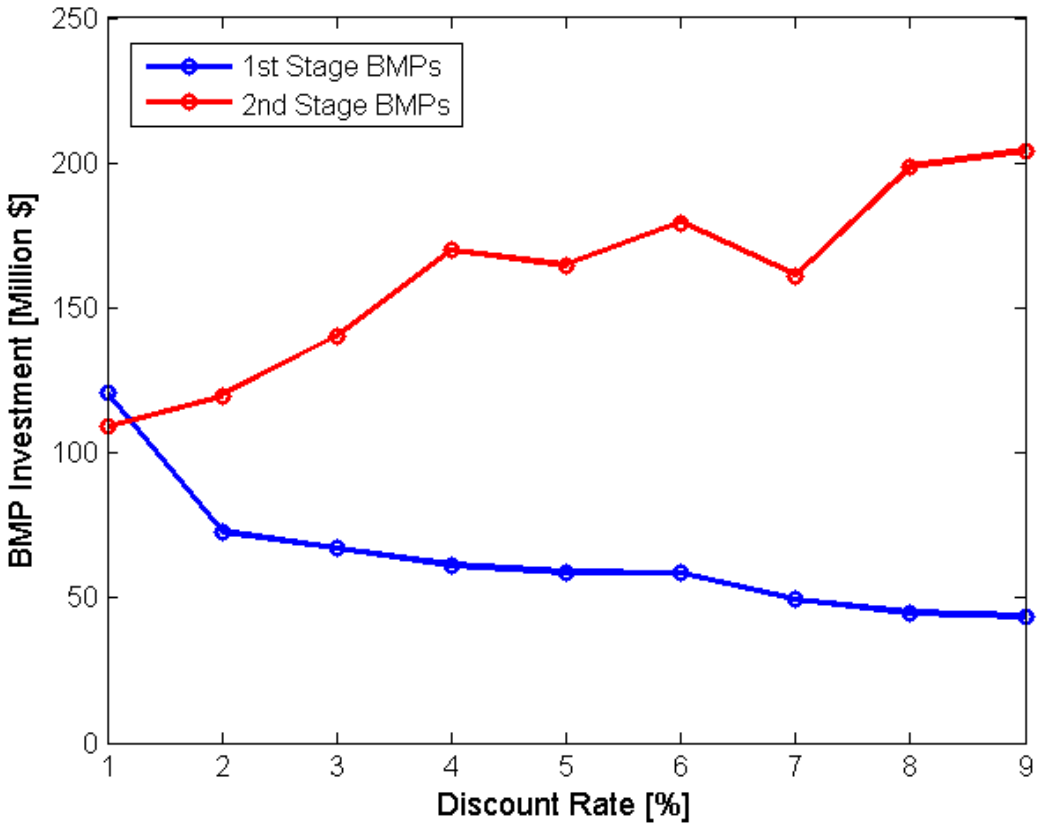

Figure 19: BMPs cost in two stages with different discount rate

.

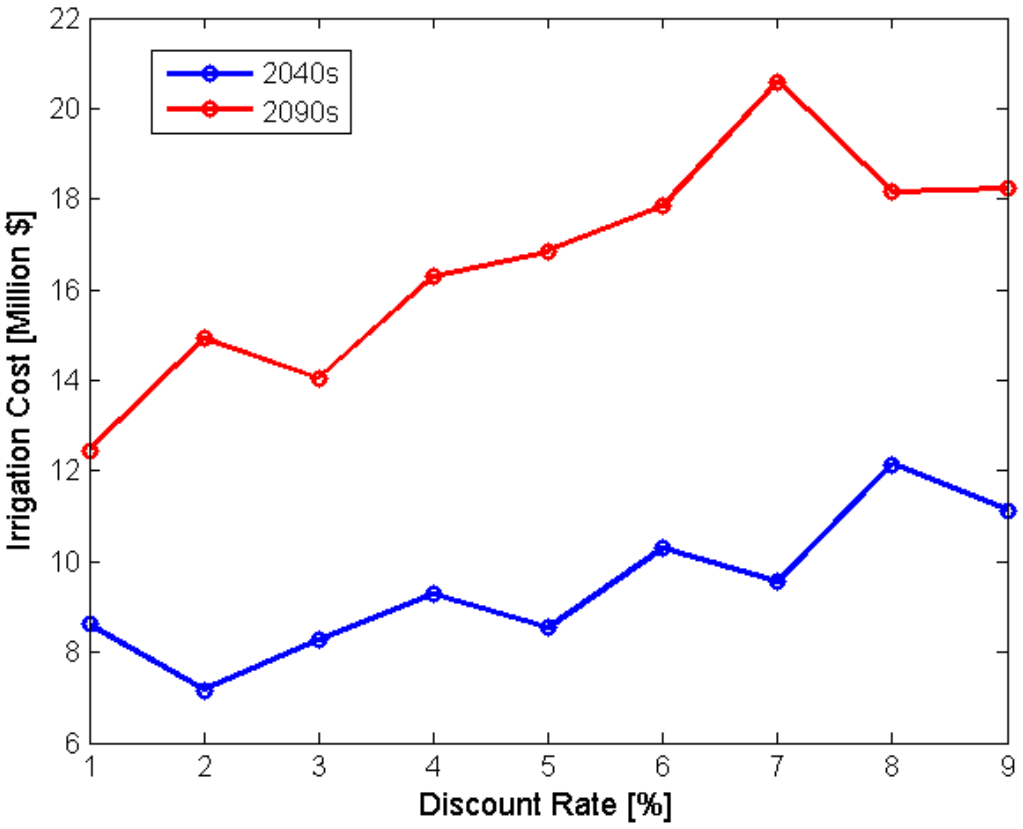

Figure 20: Irrigation operation cost in two stages change with different discount rate

.

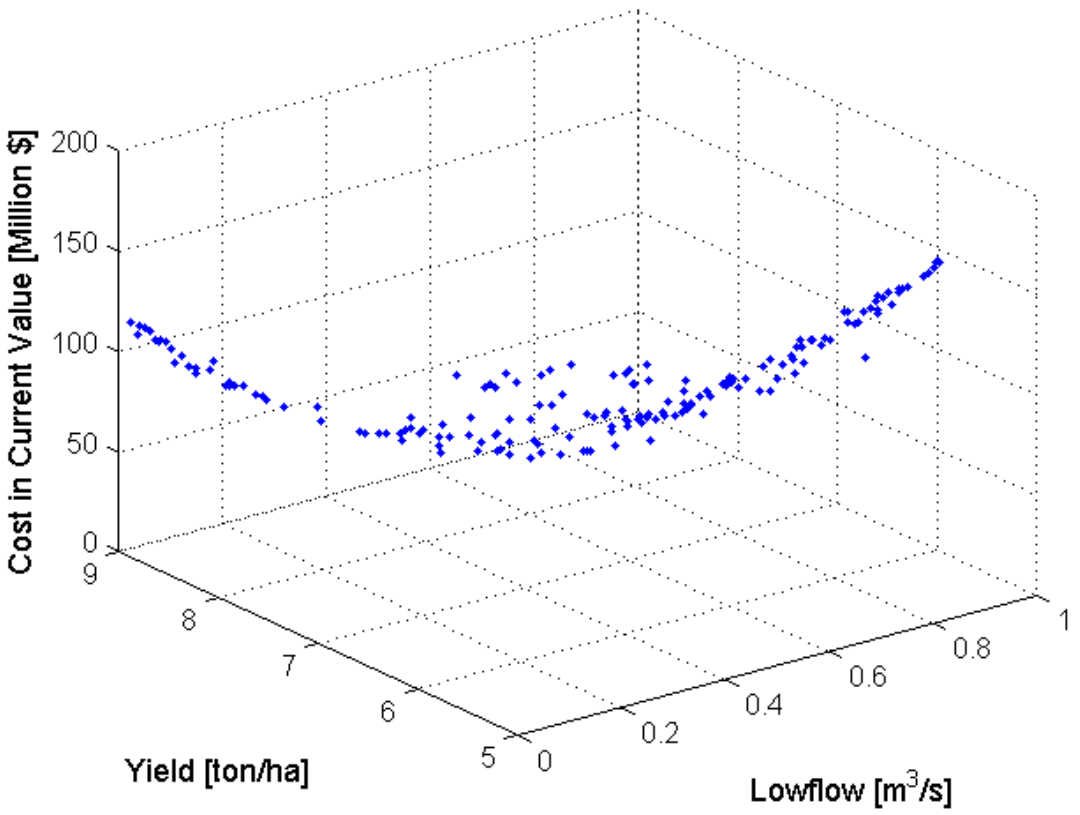

Figure 21: Pareto front of 3 objectives (cost, crop yield, 10-day-flow) when discount rate is 4%

.

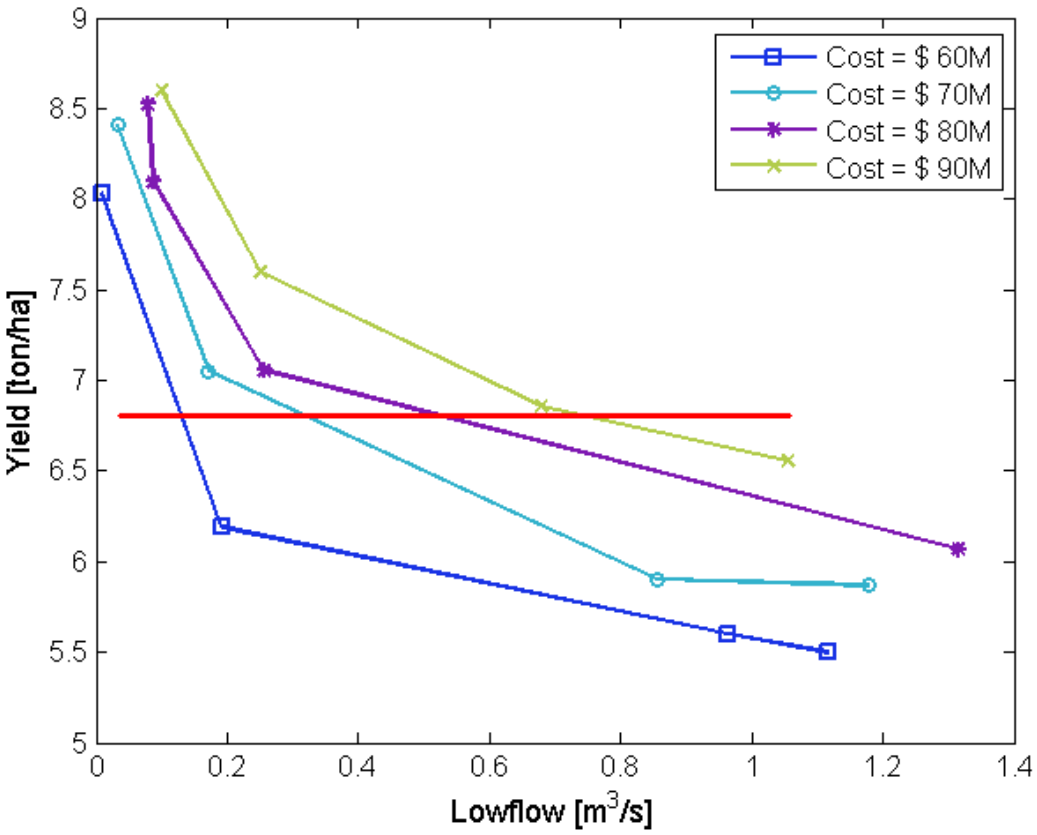

Figure 22: Trade-offs between 10-day-flow and crop yield with different cost

.

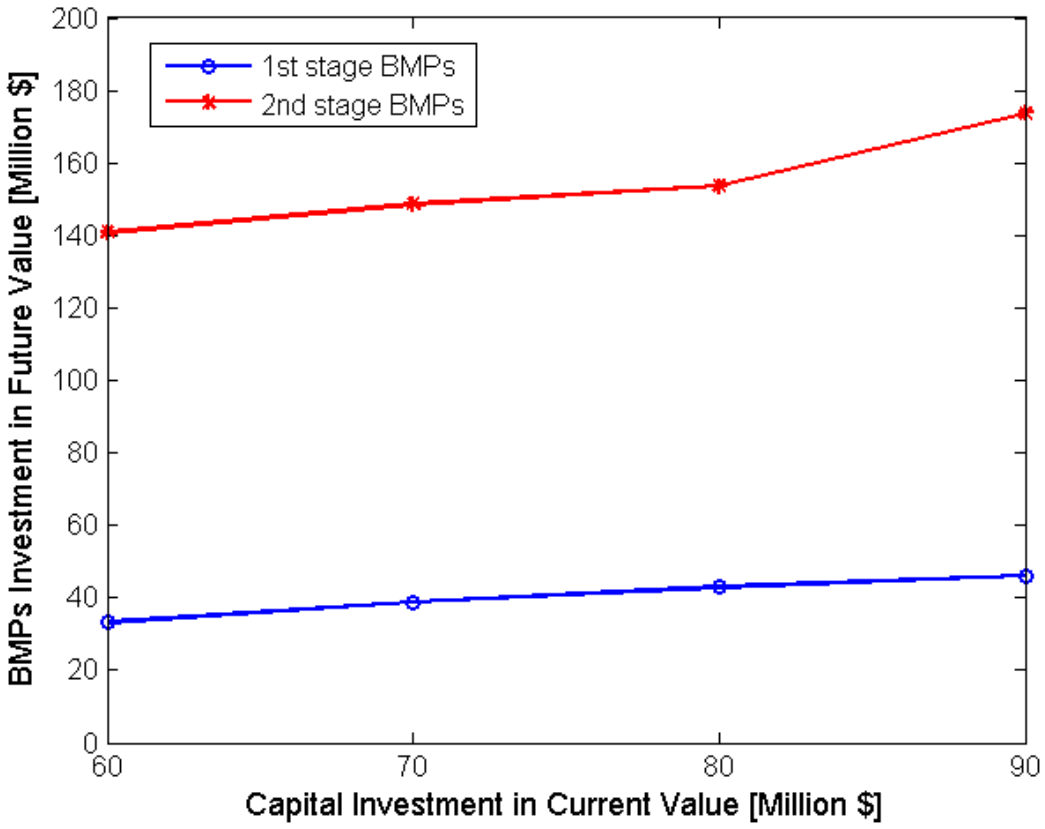

Figure 23: BMPs expansion at different stage with different investment

.

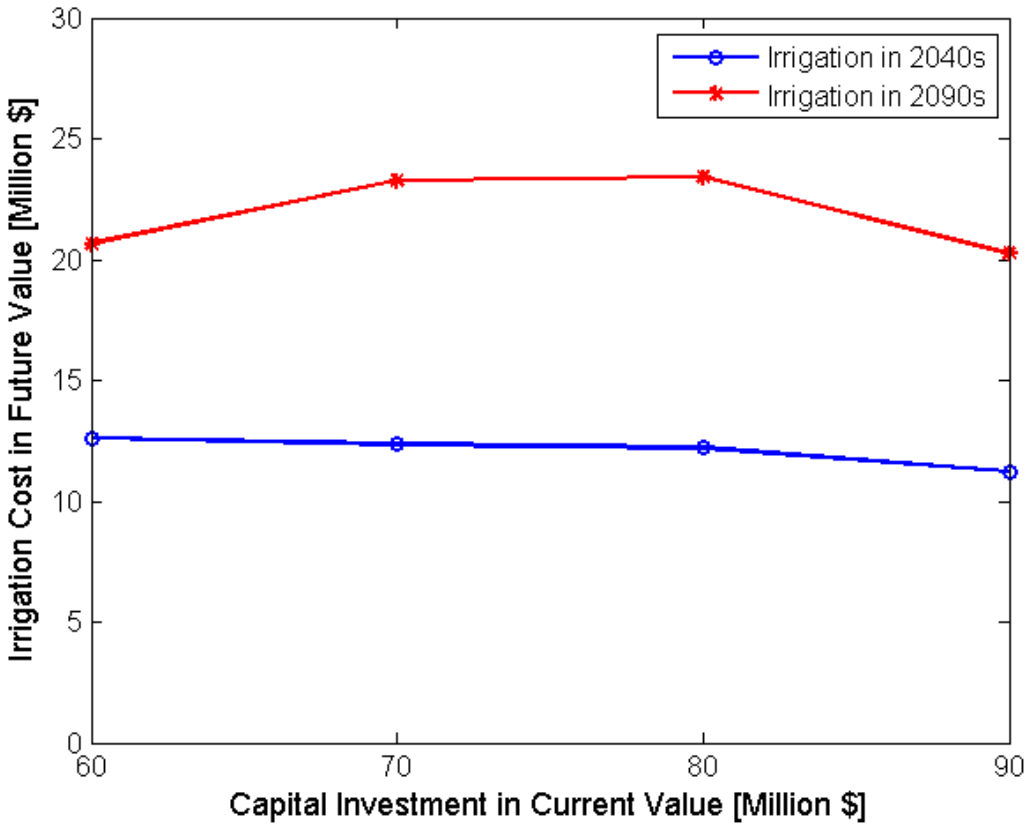

Figure 24: Irrigation operation in different periods with different investment

.

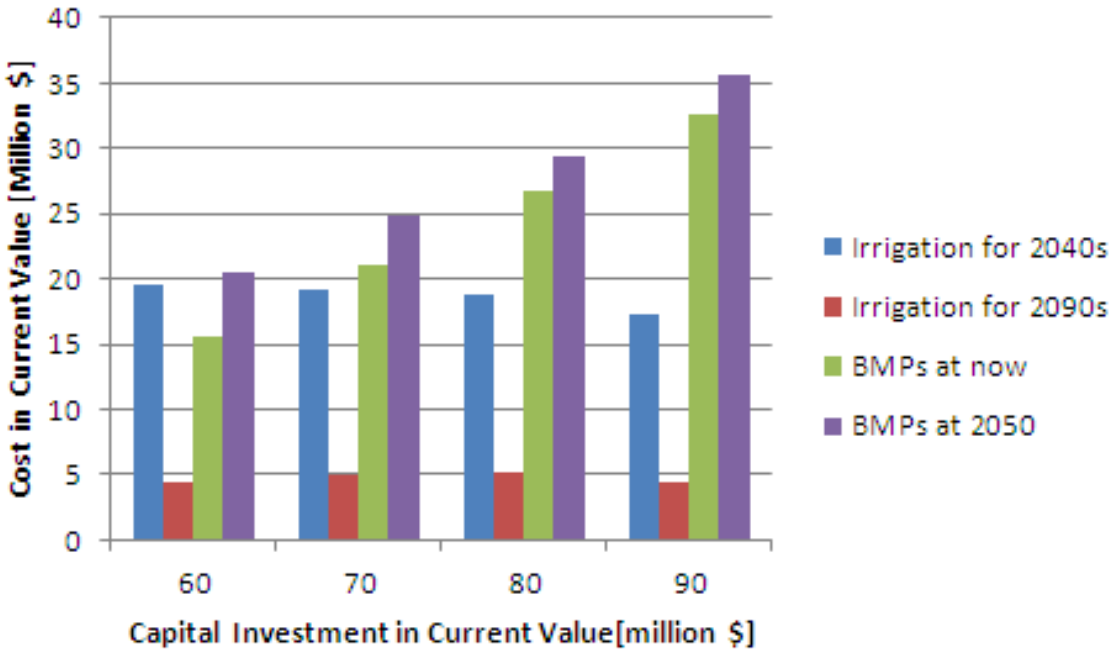

Figure 25: Investment on BMPs and irrigation at different stages with different total Investment

.

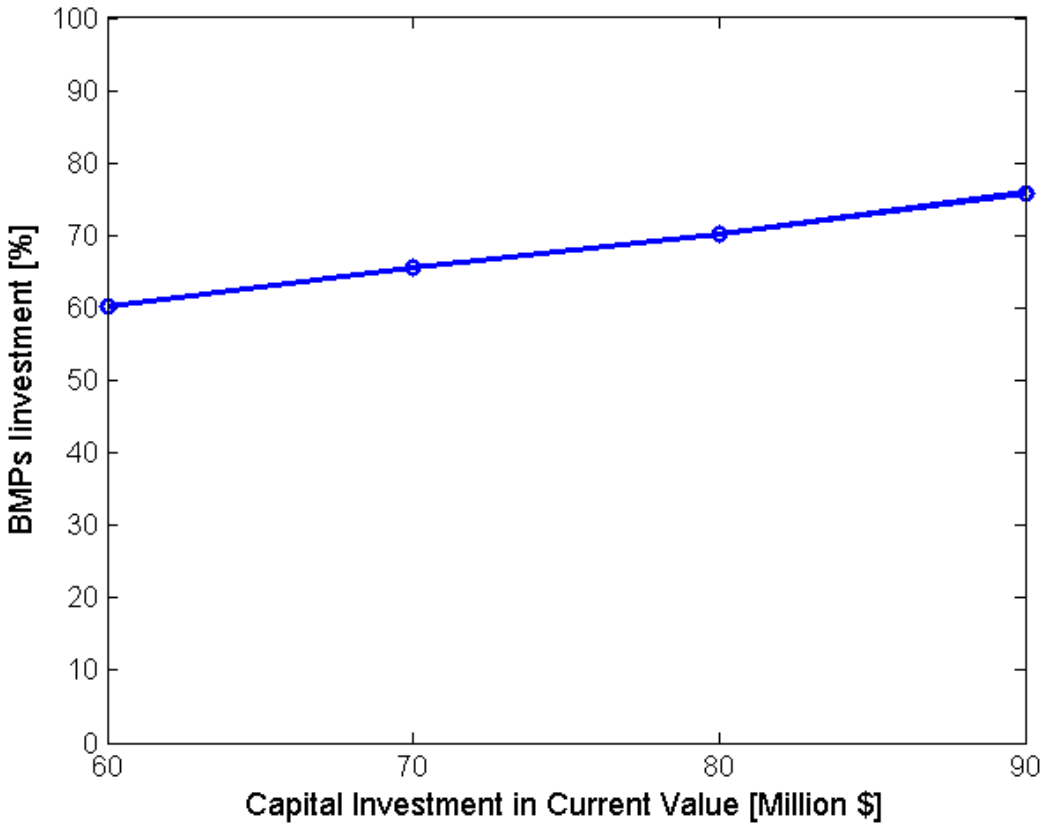

Figure 26: ratio of investment on BMPs with different total investment

.

Chapter 7

Tables

Table 1: SWAT Calibration Parameters

Parameter [Units]	Description	SWAT Default	Calibration Range	d	Calibrated value
CN [-]	Initial SCS CN II value	65	-0.2	0.2	-0.18
ALPHA[days]	Base flow alpha factor	0.048	0.01	0.2	0.0213
GW_{REVAP} [-]	Groundwater "revap" coefficient	0.02	0.01	1	0.04
GWDELAY [days]	Groundwater delay	31	5	100	43
SHALLST [m]	Initial depth of water in the shallow aquifer	0	0.01	15	12
CANMX [mm]	Maximum canopy storage	0	0.01	50	15
ESCO [-]	Soil evaporation compensation factor	0.95	0.01	1	0.03
EPCO [-]	Plant uptake compensation factor	1	0.01	1	0.728
SLSUBBSN [m]	Average slope length	121.951	1	500	345.174
K_{SAT} [mm h-1]	Saturated hydraulic conductivity	0.067-350	-0.2	0.2	0.056
EVRCH [-]	Reach evaporation adjustment factor	1	0.01	1	0.83
BIO_E	Crop solar radiation use efficiency	39	-0.2	0.2	0.035

Table 2: R-squared Values of Output from SVRs

Periods	GCM drive	scenario	R^2 in Calibration	R^2 in Validation
Baseline	PCM	Baseline	0.95	0.94
	CCSM	Baseline	0.97	0.97
	Hadley	Baseline	0.93	0.92
2040s	PCM	A1Fi	0.96	0.95
	PCM	B1	0.97	0.95
	Hadley	A2	0.94	0.92
2090s	PCM	A1Fi	0.98	0.98
	PCM	B1	0.92	0.91
	CCSM	B1	0.97	0.97
	CCSM	A1Fi	0.93	0.98
	Hadley	A2	0.96	0.96
	Hadley	B2	0.98	0.97

References

[1] S. Ahmed, A.J. King, and G. Parija. A multi-stage stochastic integer programming approach for capacity expansion under uncertainty. *Journal of Global Optimization*, 26(1):3–24, 2003.

[2] B.T. Anderson, K. Hayhoe, and X.Z. Liang. Anthropogenic-induced changes in twenty-first century summertime hydroclimatology of the northeastern us. *Climatic Change*, 99(3):403–423, 2010.

[3] M. Arabi, R.S. Govindaraju, and M.M. Hantush. Cost-effective allocation of watershed management practices using a genetic algorithm. *Water Resources Research*, 42(10):10429, 2006.

[4] JG Arnold and N. Fohrer. Swat2000: current capabilities and research opportunities in applied watershed modelling. *Hydrological Processes*, 19(3):563–572, 2005.

[5] WP Balleau. Water appropriation and transfer in a general hydrogeologic system. *Nat. Resources J.*, 28:269, 1988.

[6] A. Ben-Tal and A. Nemirovski. Robust solutions of uncertain linear programs. *Operations research letters*, 25(1):1–14, 1999.

[7] L.L. Bennett and C.W. Howe. The interstate river compact: Incentives for noncompliance. *Water Resources Research*, 34(3):485–495, 1998.

[8] J.R. Birge and F. Louveaux. *Introduction to stochastic programming*. Springer Verlag, 1997.

[9] K.S. Bracmort, BA Engel, and JR Frankenberger. Evaluation of structural best management practices 20 years after installation: Black creek watershed, indiana. *Journal of soil and water conservation*, 59(5):191–196, 2004.

[10] Oscar R. Burt, Maurice Baker, and Glenn A. Helmers. Statistical estimation of streamflow depletion from irrigation wells. *Water Resour. Res.*, 38(12):1296, 2002.

[11] A. Burton, HJ Fowler, S. Blenkinsop, and CG Kilsby. Downscaling transient climate change using a neyman-scott rectangular pulses stochastic rainfall model. *Journal of Hydrology*, 381(1-2):18–32, 2010.

[12] X. Cai and M.W. Rosegrant. Irrigation technology choices under hydrologic uncertainty: A case study from maipo river basin, chile. *Water Resources Research*, 40(4):W04103, 2004.

[13] C.C. Chang and C.J. Lin. Libsvm: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(3):27, 2011.

[14] X. Chen and L. Shu. Stream-aquifer interactions: Evaluation of depletion volume and residual effects from ground water pumping. *Ground Water*, 40(3):284–290, 2002.

[15] V.T. Chow, D.R. Maidment, and L.W. Mays. *Applied hydrology*. 1988.

[16] I.M. Chung, N.W. Kim, J. Lee, and M. Sophocleous. Assessing distributed groundwater recharge rate using integrated surface water-groundwater modelling: application to mihocheon watershed, south korea. *Hydrogeology Journal*, 18(5):1253–1264, 2010.

[17] M. P. Clark, D. E. Rupp, R. A. Woods, H. J. Tromp-van Meerveld, N. E. Peters, and J. E. Freer. Consistency between hydrological models and field observations: linking processes at the hillslope scale to hydrological responses at the watershed scale. *Hydrological Processes*, 23(2):311–319, 2009.

[18] W.D. Collins, C.M. Bitz, M.L. Blackmon, G.B. Bonan, C.S. Bretherton, J.A. Carton, P. Chang, S.C. Doney, J.J. Hack, and T.B. Henderson. The community climate system model version 3 (ccsm3). *Journal of Climate*, 19(11):2122–2143, 2006.

[19] C. Covey, K.M. AchutaRao, U. Cubasch, P. Jones, S.J. Lambert, M.E. Mann, T.J. Phillips, and K.E. Taylor. An overview of results from the coupled model intercomparison project. *Global and Planetary Change*, 37(1):103–133, 2003.

[20] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: Nsga-ii. *Evolutionary Computation, IEEE Transactions on*, 6(2):182–197, 2002.

[21] J. Dudhia, D. Gill, K. Manning, W. Wang, and C. Bruyere. Psu/ncar mesoscale modeling system tutorial class notes and users guide: Mm5 modeling system version 3. *National Center for Atmospheric Research*, 2005.

[22] J. Dupacov, G. Consigli, and S.W. Wallace. Scenarios for multistage stochastic programs. *Annals of operations research*, 100(1):25–53, 2000.

[23] National Research Council . Committee on Estimating, Communicating Uncertainty in Weather, Climate Forecasts, National Research Council . Board on Atmospheric Sciences, and National Academies Press. *Completing the forecast: characterizing and communicating uncertainty for better decisions using weather and climate forecasts*. National Academies Press, 2006. (US).

[24] Jan H. Fleckenstein, Stefan Krause, David M. Hannah, and Fulvio Boano. Groundwater-surface water interactions: New methods and models to improve understanding of processes and dynamics. *Advances in Water Resources*, 33(11):1291–1295, 2010.

[25] P.W. Gassman, M.R. Reyes, C.H. Green, and J.G. Arnold. The soil and water assessment tool: Historical development, applications, and future research directions. 2007.

[26] MW Gitau, T.L. Veith, and WJ Gburek. Farm-level optimization of bmp placement for cost-effective pollution reduction. 2004.

[27] Pascal Goderniaux, Serge Brouyre, Stephen Blenkinsop, Aidan Burton, Hayley J. Fowler, Philippe Orban, and Alain Dassargues. Modeling climate change impacts on groundwater resources using transient stochastic climatic scenarios. *Water Resour. Res.*, 47(12):W12516, 2011.

[28] G.A. Griffiths and B. Clausen. Streamflow recession in basins with multiple water storages. *Journal of Hydrology*, 190(1):60–74, 1997.

[29] N. Hibiki. A hybrid simulation/tree stochastic optimization model for dynamic asset allocation. *Asset and Liability Management Tools: A Handbook for Best Practice*, pages 269–294, 2003.

[30] H.S.J. Hill, J.W. Mjelde, H.A. Love, D.J. Rubas, S.W. Fuller, W. Rosenthal, and G. Hammer. Implications of seasonal climate forecasts on world wheat trade: a stochastic, dynamic analysis. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, 52(3):289–312, 2004.

[31] GH Huang and DP Loucks. An inexact two-stage stochastic programming model for water resources management under uncertainty. *CIVIL ENGINEERING SYSTEMS*, 17(2):95–118, 2000.

[32] Yanbing Jia and Teresa B. Culver. Robust optimization for total maximum daily load allocations. *Water Resour. Res.*, 42(2):W02412, 2006.

[33] P. Kaini, K. Artita, and JW Nicklow. Evaluating optimal detention pond locations at a watershed scale. pages 1–8. ASCE.

[34] P. Kaini, K. Artita, and J.W. Nicklow. Optimizing structural best management practices using swat and genetic algorithm to improve water quality goals. *Water Resources Management*, pages 1–19, 2012.

[35] N.W. Kim, I.M. Chung, Y.S. Won, and J.G. Arnold. Development and application of the integrated swatmodflow model. *Journal of Hydrology*, 356(1):1–16, 2008.

[36] R. Laurent and X. Cai. A maximum entropy method for combining aogcms for regional intra-year climate change assessment. *Climatic Change*, 82(3):411–435, 2007.

[37] YP Li and GH Huang. Two-stage planning for sustainable water-quality management under uncertainty. *Journal of environmental management*, 90(8):2402–2413, 2009.

[38] YP Li, GH Huang, and SL Nie. Water resources management and planning under uncertainty: an inexact multistage joint-probabilistic programming method. *Water Resources Management*, 23(12):2515–2538, 2009.

[39] X.Z. Liang, K.E. Kunkel, G.A. Meehl, R.G. Jones, and J.X.L. Wang. Regional climate models downscaling analysis of general circulation models present climate biases propagation into future change projections. *Geophysical research letters*, 35(8):L08709, 2008.

[40] X.Z. Liang, K.E. Kunkel, and A.N. Samel. Development of a regional climate model for us midwest applications. part i: Sensitivity to buffer zone treatment. *Journal of Climate*, 14(23):4363–4378, 2001.

[41] X.Z. Liang, L. Li, K.E. Kunkel, M. Ting, and J.X.L. Wang. Regional climate model simulation of us precipitation during 1982-2002. part i: Annual cycle. *Journal of Climate*, 17(18):3510–3529, 2004.

[42] X.Z. Liang, J. Pan, J. Zhu, K.E. Kunkel, J.X.L. Wang, and A. Dai. Regional climate model downscaling of the us summer climate and future change. *J. Geophys. Res*, 111:D10108, 2006.

[43] E.T. Maillet. *Essais d'hydraulique souterraine et fluviale*. A. Hermann, 1905.

[44] G.F. Marques, J.R. Lund, and R.E. Howitt. Modeling irrigated agricultural production and water use decisions under water supply uncertainty. *Water Resources Research*, 41(8):W08423, 2005.

[45] J.W. Massman, Washington State Transportation Commission. Planning, Capital Program Management, and Washington . Dept. of Transportation. A design manual for sizing infiltration ponds. Technical report, Washington State Department of Transportation, 2003. (State).

[46] J.M. Mulvey, R.J. Vanderbei, and S.A. Zenios. Robust optimization of large-scale systems. *Operations research*, pages 264–281, 1995.

[47] N. Nakicenovic, J. Alcamo, G. Davis, B. de Vries, J. Fenhann, S. Gaffin, K. Gregory, A. Grubler, T.Y. Jung, and T. Kram. Special report on emissions scenarios: a special report of working group iii of the intergovernmental panel on climate change. Technical report, Pacific Northwest National Laboratory, Richland, WA (US), Environmental Molecular Sciences Laboratory (US), 2000.

[48] J.E. Nash and JV Sutcliffe. River flow forecasting through conceptual models part ia discussion of principles. *Journal of Hydrology*, 10(3):282–290, 1970.

[49] R.K. Pachauri. *Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, volume 446. IPCC, 2007.

[50] R.N. Palmer, S.L. Kutzing, and A.C. Steinemann. Developing drought triggers and drought responses: an application in georgia. pages 19–22.

[51] VD Pope, ML Gallani, PR Rowntree, and RA Stratton. The impact of new physical parametrizations in the hadley centre climate model: Hadam3. *Climate Dynamics*, 16(2):123–146, 2000.

[52] H. Pranevicius and K. utiene. Scenario tree generation by clustering the simulated data paths. pages 203–208.

[53] T. Ross, N. Lott, and National Climatic Data Center. *A climatology of 1980-2003 extreme weather and climate events*. US Department of Commerece, National Ocanic and Atmospheric Administration, National Environmental Satellite Data and Information Service, National Climatic Data Center, 2003. (US).

[54] A. Saleh and B. Du. Evaluation of swat and hspf within basins program for the upper north bosque river watershed in central texas. *Transactions of the ASAE*, 47(4):1039–1049, 2004.

[55] C. Santhi, JG Arnold, JR Williams, LM Hauck, and WA Dugas. Application of a watershed model to evaluate management effects on point and nonpoint source pollution. *Transactions of the ASAE*, 44(6):1559–1570, 2001.

[56] A.J. Smola and B. Schlkopf. A tutorial on support vector regression. *Statistics and computing*, 14(3):199–222, 2004.

[57] M. Sophocleous. On the elusive concept of safe yield and the response of interconnected stream-aquifer systems to development. *BULLETIN-KANSAS GEOLOGICAL SURVEY*, pages 61–86, 1998.

[58] M. Sophocleous, A. Koussis, JL Martin, and SP Perkins. Evaluation of simplified stream-aquifer depletion models for water rights administration. *Ground Water*, 33(4):579–588, 1995.

[59] M. Sophocleous and S.P. Perkins. Methodology and application of combined watershed and ground-water models in kansas. *Journal of Hydrology*, 236(3):185–201, 2000.

[60] C.P. Spalding and R. Khaleel. An evaluation of analytical solutions to estimate drawdowns and stream depletions by wells. *Water Resources Research*, 27(4):597–609, 1991.

[61] J. Szilagyi. Streamflow depletion investigations in the republican river basin: Colorado, nebraska, and kansas. *Journal of Environmental Systems*, 27(3):251–263, 1999.

[62] J. Szilagyi. Identifying cause of declining flows in the republican river. *Journal of water resources planning and management*, 127(4):244–253, 2001.

[63] LM Tallaksen. A review of baseflow recession analysis. *Journal of Hydrology*, 165(1):349–370, 1995.

[64] V. Vapnik. Statistical learning theory. 1998, 1998.

[65] WM Washington, JW Weatherly, GA Meehl, AJ Semtner Jr, TW Bettge, AP Craig, WG Strand Jr, J. Arblaster, VB Wayland, and R. James. Parallel climate model (pcm) control and transient simulations. *Climate Dynamics*, 16(10):755–774, 2000.

[66] D.W. Watkins Jr. Finding robust solutions to water resources problems. *Journal of water resources planning and management*, 123:49, 1997.

[67] D.W. Watkins Jr, D.C. McKinney, L.S. Lasdon, S.S. Nielsen, and Q.W. Martin. A scenario-based stochastic programming model for water supplies from the highland lakes. *International Transactions in Operational Research*, 7(3):211–230, 2000.

[68] W.J. Werick and W. Whipple Jr. National study of water management during drought: managing water for drought. Technical report, DTIC Document, 1994.

[69] O. Wilchfort. Shortage management modeling for urban water supply systems. *Journal of water resources planning and management*, 123:250, 1997.

[70] D.A. Wilhite. Combating drought through preparedness. volume 26, pages 275–285. Wiley Online Library.

[71] D.A. Wilhite, M.J. Hayes, C. Knutson, and K.H. Smith. Planning for drought: Moving from crisis to risk management1. *JAWRA Journal of the American Water Resources Association*, 36(4):697–710, 2000.

[72] D.A. Wilhite, M.J. Hayes, C. Knutson, and K.H. Smith. Planning for drought: Moving from crisis to risk management1. *JAWRA Journal of the American Water Resources Association*, 36(4):697–710, 2000.

[73] T.C. Winter. *Ground water and surface water: a single resource*. DIANE Publishing, 1999.

[74] H. Wittenberg. Baseflow recession and recharge as nonlinear storage processes. *Hydrological Processes*, 13(5):715–726, 1999.

[75] H. Wittenberg. Baseflow recession and recharge as nonlinear storage processes. *Hydrological Processes*, 13(5):715–726, 1999.

[76] Z. Yi, P.A. Heng, and A.W.C. Fu. Estimate of exponential convergence rate and exponential stability for neural networks. *Neural Networks, IEEE Transactions on*, 10(6):1487–1493, 1999.

[77] L.Y. Yu, X.D. Ji, and S.Y. Wang. Stochastic programming models in financial optimization: A survey. 2003.